HAPLOTYPE-BASED PROSTATE CANCER ASSOCIATION STUDY IN AFRO-CARIBBEANS OF TOBAGO

by

You Li

B.S, Biology, Zhejiang University, China, 2003

Submitted to the Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2010

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This dissertation was presented

by

You Li

It was defended on

March 24, 2010

and approved by

Dissertation Advisor:
Robert E. Ferrell, PhD, Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Jeremy Martinson, PhD, Assistant Professor
Department of Infectious Diseases and Microbiology
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Clareann Bunker, PhD, Associate Professor
Department of Epidemiology
Graduate School of Public Health
University of Pittsburgh

Committee Member:
Eleanor Feingold, PhD, Professor
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Robert E. Ferrell, PhD

HAPLOTYPE-BASED PROSTATE CANCER ASSOCIATION STUDY IN AFRO-CARIBBEANS OF TOBAGO

You Li, PhD

University of Pittsburgh, 2010

Prostate cancer (PC) remains a significant public health concern for men throughout the world. Identification of the environmental and genetic factors that predispose to PC for prevention or early intervention is a significant public health concern. PC is more prevalent in males of African descent living in the western hemisphere. Also, recent studies have shown high rates of PC in Afro-Caribbean populations. Men in the 40-79 year age range from the Caribbean island of Tobago have a 3 fold higher risk of developing prostate cancer than do Caucasians. We performed two candidate gene-based association studies to explore the genetic determinants for high risk of PC in Tobago population.

In the first study, we performed an extensive study to identify sequence variation in the DC-SIGN gene, a candidate gene for PC, in the Tobago population, and carried out a case/control association study of DC-SIGN polymorphisms and PC susceptibility in the high-risk population of Tobago. We found a unique haplotype in the 5' proximal promoter of DC-SIGN that is associated with PC in Tobago. We further extended the association study to single gene polymorphisms likely to be in linkage disequilibrium with DC-SIGN. We also found that a single nucleotide polymorphism, rs4804806, which is located in the intragenic region between the DC-SIGN and DCSIGN-R genes (4228bp upstream of DC-SIGN), is significantly associated

with elevated risk of PC in Tobago population. The function of rs4804806, and whether the risk associated with it is due to linkage disequilibrium between it and other causal variation remains to be explored.

In the second study, we examined whether genetic variation in the androgen receptor (AR) and PSA genes was associated with risk of prostate cancer or with serum PSA levels in the Tobago population. Previous studies have proposed that the rs266882 G/A polymorphism, located in the androgen response element upstream of the PSA gene and the $(CAG)_n$ repeat polymorphism in DNA binding domain of the AR are associated with risk of PC and serum PSA levels, and those two genetic variants show evidence of genetic interaction. In order to examine whether those two genetic variants affect PC risk or PSA level in Tobago population, an association study was carried out in 167 PC cases and 320 controls. Association analysis with PC was conducted among all cases and controls for each locus individually, and also for both loci together, to test for interaction between them in determining risk of PC or PSA levels. Regression analysis was carried out for each locus and combined with serum PSA level in 320 controls only. Results from this study do not provide any association evidence with PC risk or PSA level for PSA rs266882 G/A genotypes or AR $(CAG)_n$ repeat length genotypes.

The results from these two studies lend support to exploring genetic causes and prevention of prostate cancer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

At the beginning of my dissertation, I would like to take this opportunity here to express my sincere gratitude to the following.

My greatest appreciations go to my committee members Dr. Robert E. Ferrell, Dr. Jeremy Martinson, Dr. Clareann Bunker and Dr. Eleanor Feingold. I would have never fulfilled this project without all your invaluable advice and kindly help. I would like to thank my advisor Dr. Ferrell and Dr. Martinson for all your guidance through my work. You helped on every aspect of my research. I would like to thank my thesis committee Dr. Bunker for your countless help on experiment design and providing me all the samples for this project. I would also like to thank my committee member Dr. Feingold for her help with statistical analysis in my study.

I would specially like to thank my parents and my wife, Ms. Dan Wang. Your always love and encouragement means most to me, and I wouldn't have made it through some of those days without you.

Many thanks go to all my lab mates, Elizabeth, Mark, Nancy, Chris, Xiaojing and Becky, thanks for all your help!

I would like to thank all members in department of Human Genetics, especially Jeanette Norbut, Michele Lavalley and Gloria Duval, for being nice and helpful to my study here, Chris Huziak and Terry Bates for computer technical supports.

Finally, thank you all my friends!

1.0 CHAPTER ONE: INTRODUCTION

1.1 Epidemiology of Prostate Cancer

Prostate cancer (PC) is highly prevalent, being the most common non-cutaneous cancer, and the second most common cause of cancer death in men. Globally an estimated 782,600 new cases and 254,000 deaths caused by PC occurred in 2007. (Crawford 2009) In the United States, PC accounts for about 25% of all cancers diagnosed in men in 2008. In 2008, the 5-year survival rate for localized or regional PC was very close to 100%. However when it has already metastasized by time of diagnosis, the 5-year survival rate has been reported to be as low as 32%. (Jemal, Siegel et al. 2008) With the introduction of prostate specific antigen (PSA) testing, and digital rectal examination (DRE) screening, more and more prostate cancer cases are being diagnosed at earlier stage. The only firmly established risk factors for PC are age, ethnicity and family history of PC.

PC is a disease strongly associated with aging. The probability of developing PC increases from 0.005% among individuals aged <39 years to 2.2% among those aged 40-59 years and 13.7% for Caucasian men, aged over 60-79%. (Crawford 2003) Individuals of African descent have the highest rates of PC in the world. In the United States, the incidence among African Americans is nearly 60% higher than among Caucasians. Moreover, the mortality rate for African Americans is about twice as high as for whites, 3.3 times higher than for Hispanics and 5 times higher than for Asians. (Crawford 2003)

In a recent study from the Caribbean nation of Trinidad and Tobago, the prevalence of screening-detected prostate cancer was three-fold higher among Tobago men of African descent compared to US Caucasian men. (Bunker, Patrick et al. 2002; Bunker, Patrick et al. 2002) This variation between ethnic groups may be related to differences in genetic background, environment factors (e.g. diet, environmental exposures) and physiologic status (e.g. hormone levels).

Prostate cancer is highly heritable. The risk of developing PC doubles for men who have a father or brother affected by PC, and epidemiologic studies indicate that men with a positive family history are diagnosed on average 6-7 years earlier than those without an affected first-degree relative. These studies estimate that about 10% of all PC cases and up to 40% of those occurring at <55 years of age may have a hereditary basis. (Steinberg, Carter et al. 1990; Bratt 2002) These recurrence rates are estimated primarily based on Caucasian men, while the role of heredity in populations of African descent is not well characterized.

Recently, some additional factors have been reported to be associated with risk of PC incidence. For example, a higher intake of total fat is associated with higher risk of PC, and total intake of soybean products and tomato products are reported to be protective factors associated with lower risk of PC. (Giovannucci, Rimm et al. 1993; Whittemore, Kolonel et al. 1995) However, these results still remain to be replicated.

More and more evidence suggests that chronic inflammation is a risk factor in PC. Firstly, chronic inflammation can cause clinical prostatitis and prostatic proliferative inflammatory atrophy (PIA), which has been suggested to be associated with PC and a potential PC precursor lesion. Secondly, studies have shown that several sexually transmitted infections

(STIs) are associated with PC (e.g. human papilloma virus strains HPV-16 and HPV-18).(Tu, Jacobs et al. 1994; Serth, Panitz et al. 1999) Thirdly, several innate immune response related genes have been found to be associated with PC, such as the endo-ribonuclease gene (RNASEL) (Casey, Neville et al. 2002; Wang, McDonnell et al. 2002), macrophage scavenger receptor (MSR1) (Xu, Zheng et al. 2002; Xu, Zheng et al. 2003; Lindmark, Jonsson et al. 2004), and glutathione S-transferase (GSTP1) (Nelson, Kidd et al. 2001; Jeronimo, Varzim et al. 2002). These immune related genes play important roles in regulating pathogen infection and the inflammatory response.

Recently, Hoffman et al (2004) and McDonald et al (2009) have suggested that human herpesvirus 8 (HHV8) may play a role in PC because of the elevated prevalence of HHV8 seropositivity among PC cases in Tobago and among women in Tobago. (Hoffman, Bunker et al. 2004; McDonald, Ragin et al. 2009)

HHV-8, also known as Kaposi's sarcoma associated herpesvirus, is the most recently discovered human herpesvirus. (Chang, Cesarman et al. 1994; Rivas, Thlick et al. 2001) It is a member of the lymphotrophic (gamma) herpesviruses. (Rivas, Thlick et al. 2001) HHV8 is known to cause Kaposi's sarcoma, primary effusion lymphoma, and is reported to be associated with several other cancers. (Moore and Chang 2001; Rivas, Thlick et al. 2001; Bollen, Polstra et al. 2003) DC-SIGN serves as a receptor for HHV8. (Rappocciolo, Jenkins et al. 2006) Based on these findings, we hypothesized that DC-SIGN may play a role in regulating infections and prostatic inflammation and therefore contribute to the elevated risk of PC in Tobago.

1.2 Genetics of Prostate Cancer

Prostate cancer is one of the most heritable late onset cancers. Studies have suggested that around 10% of all PC cases, and up to 40% of those occurring before 55 years of age, have a hereditary basis. (Carter, Beaty et al. 1992; Bratt, Damber et al. 2002) However, these studies are carried out primarily in Caucasian populations, and the heritability of PC among other ethnic groups is not well characterized. There have been numerous genome wide and targeted linkage studies of PC, and implicating a large number of chromosome regions, including 1q, 2p, 6p, 8q, 16p, 17p, 19q, 20 and the X chromosome, but these have largely failed to produce reproducible results. Positional cloning has successfully identified RNASEL (HPC1, 1q24-25), MSR1 (8p32) and ELAC2 (HPC2, 17p2) as genes located within linked intervals and having mutations in prostate cancer families or patients. (Tavtigian, Simard et al. 2001; Carpten, Nupponen et al. 2002; Xu, Zheng et al. 2002) In the pre-genome wide association study (GWAS) era, numerous single gene associations were reported between PC and a wide spectrum of genes, and these studies were either replicated or not in subsequent studies, making any conclusion about the genetic basis of PC questionable. (Vastag 2002; Freedman, Haiman et al. 2006; Gudmundsson, Sulem et al. 2007; Gudmundsson, Sulem et al. 2007; Yeager, Orr et al. 2007; Witte 2009)

With the advent of GWAS studies, there have been several studies of PC reported in various populations, and significant association has been reported for SNP on almost every chromosome. It has been concluded from these studies that PC is a complex disease with multiple individual genes, each with small effects contributing to risk. Few of these studies, both linkage and association, have included a sufficient number of individuals of African descent to identify specific genes or generalize the results of studies in Caucasian populations to populations of African ancestry. Interestingly none of the regions previously reported from

linkage studies or biological candidate studies are consistently found by GWAS. (Amundadottir, Sulem et al. 2006; Gudmundsson, Sulem et al. 2007; Yeager, Orr et al. 2007; Eeles, Kote-Jarai et al. 2008; Thomas, Jacobs et al. 2008) Whether this indicates that these play a minor role in determining susceptibility in the general population, or represent false positive findings is not known.

PC is a testosterone dependent tumor, at least at the early stage, leading investigations to examine genes involved in hormone metabolism. Among the most intensively studied biological candidate genes for PC is the transactivation domain encoded by exon 1 of the androgen receptor (AR) gene, which has 2 different tri-nucleotide repeat variants (CAG and GGC). Some studies have suggested that a shorter (CAG)n repeat length is associated with increased PC risk, while other studies have failed to confirm such association. (Xue, Irvine et al. 2000; Chang, Zheng et al. 2002; Gsur, Preyer et al. 2002; Medeiros, Morais et al. 2002; Salinas, Austin et al. 2005) The consensus is that the variation is associated with a small but measureable increase in risk of PC. Another candidate gene is SRD5A2, which encodes the 5α-reductase type 2 that catalyzes the conversion of testosterone to the more active dihydrotestosterone. The Ala49Thr mutation in SRD5A2 increases its catalytic activity and could therefore increase PC risk. (Makridakis, Ross et al. 1999) The results remain equivocal. (Li, Coates et al. ; Ntais, Polycarpou et al. 2003)

Genes involved in immune response are also promising targets for PC candidate-gene studies, because those genes play important role in infections and chronic inflammation, both of which are important PC risk factors. Recent studies have shown that sequence variation in genes related to the early immune response, such as the endo-ribonuclease gene RNASEL, macrophage scavenger receptor (MSR), glutathione S-transferase (GSTP1) increase the risk of PC. (Palapattu, Sutcliffe et al. 2005)

1.3 Tobago Population

Tobago is the smaller of the two main islands that make up the Republic of Trinidad and Tobago. It is located in the southern Caribbean Sea, northeast of the island of Trinidad and south of Grenada. Tobago is a land area of 300 km^2, and is approximately 42km long and 10km wide. (Fig. 1) The population is around 60,000. While Trinidad is multiethnic, the Tobago population is overwhelmingly Afro-Tobagonian. The population of Tobago was reported to be 92 percent black, 4.5 percent mixed, 2 percent East Indian, 0.4 percent white, and 1 percent other in the 1990 census. This distribution is very different from the 38 percent black, 20 percent mixed, 42 percent East Indian, and 0.3 percent white/other reported in the total Trinidad & Tobago population. (Central Statistical Office of Tobago 1993; Boyd-Patrick HA 1997) The older population of the island appears to be a stable population of descendents of West African forbears who arrived around 200 years ago. They are thought to share considerable West African Ancestry with African Americans. The Tobago population of non-African descent has been very small throughout this time period, and genetic admixture appears to be relatively low. (Miljkovic-Gacic, Ferrell et al. 2005) This low admixture rate makes it very suitable population for genetic research.

Men in the 40-79 year age range from the Caribbean island of Tobago have a 3-4 times higher risk of developing prostate cancer than do Caucasians.(Bunker, Patrick et al. 2002) And elevated human herpesvirus 8 (HHV-8) seroprevalence has been reported in among Tobago prostate cancer cases. (Hoffman, Bunker et al. 2004) Therefore, we hypothesize that genetic variation in the HHV-8 immune receptor DC-SIGN is associated with elevated HHV-8 seroprevalence and risk of prostate cancer in men of Tobago in our first candidate-gene based PC study.

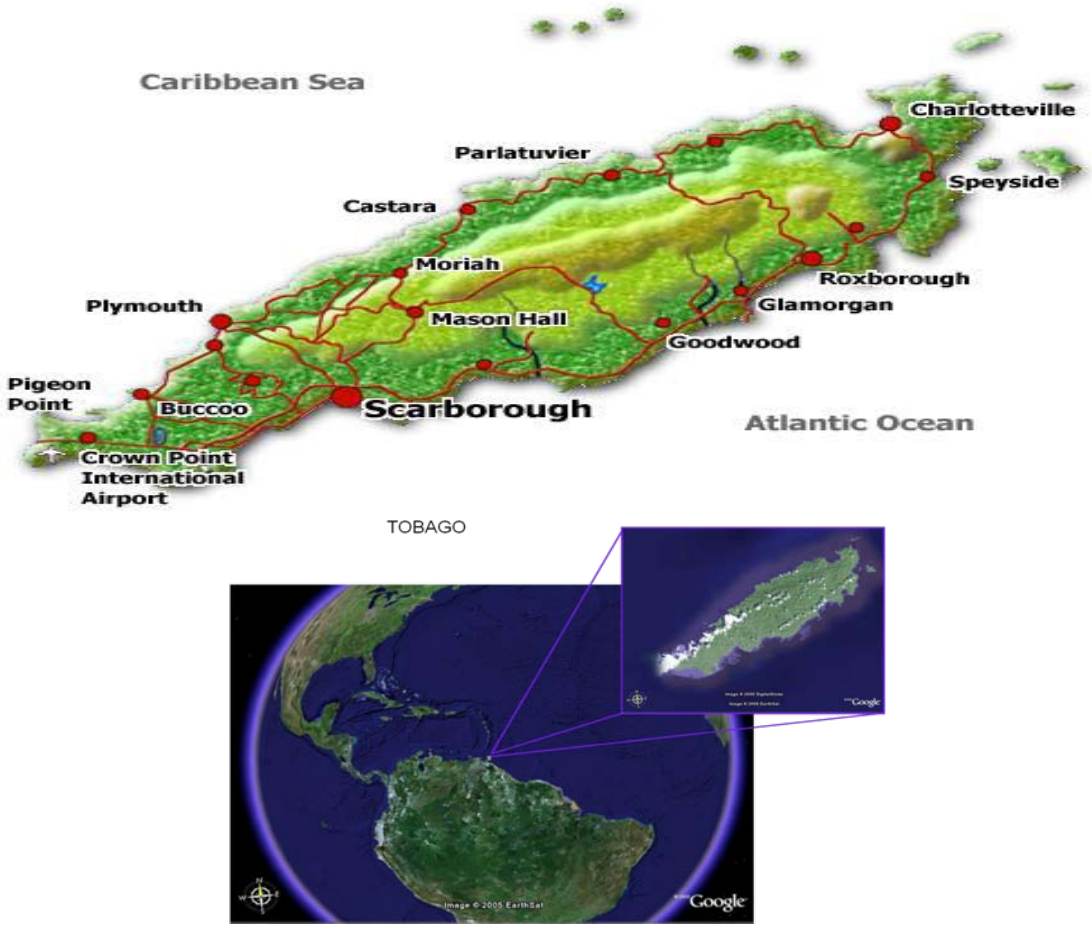

Figure 1: Geographic information on Tobago

Geographic view of Tobago generated by Google Earth software. (earth.google.com)

2.0 CHAPTER TWO: PROSTATE CANCER ASSOCIATION STUDY OF DC-SIGN GENE IN THE TOBAGO POPULATION

2.1 Background and Introduction to DC-SIGN

Dendritic cell-specific ICAM-3-grabbing non-integrin (DC-SIGN; CD209) is a type II C-type lectin that functions as an adhesion receptor and mediates viral binding and internalization. It is expressed in dendritic cells (DCs), macrophages, and B cells, and is exploited by many pathogens, such as HIV-1, Ebola virus, and HHV-8, to gain entry to the cell. (Geijtenbeek, Kwon et al. 2000; Alvarez, Lasala et al. 2002; Rappocciolo, Jenkins et al. 2006) Interestingly, epidemiological data has suggested that PC cases in Tobago also have an elevated rate of viral infection including HHV-8. (Hoffman, Bunker et al. 2004) Therefore, genetic variants of DC-SIGN may regulate viral infection, and further contribute to the high risk of PC in Tobago population.

The DC-SIGN gene (CD209) is located on human chromosome 19p13, in a multigene cluster that also contains genes encoding the lectins DC-SIGNR (CD209L), FceRII(CD23) and LSECtin (CLEC4G). DC-SIGN and DC-SIGNR are 73% identical at the amino acid level and both genes are heavily alternately spliced, giving rise to many isoforms with a variety of localizations. (Soilleux, Barten et al. 2000)

DC-SIGN is a small type II trans-membrane C-type lectin consisting of 404 amino acids with three distinct domains. (Geijtenbeek, Torensma et al. 2000) The N-terminal cytoplasm domain of 40 amino acid residues contains two putative internalization motifs (one di-leucine motif and a tri-acidic cluster motif), and is critical for antigen internalization. The transmembrane domain consists of 7 complete and one incomplete repeats of nearly identical sequences, each repeat encoding 23 amino acids, and is important for DC-SIGN oligomerization, carbohydrate recognition and normal DC function. (Geijtenbeek, Torensma et al. 2000; Feinberg, Guo et al. 2005; Snyder, Ford et al. 2005) It has recently been reported that the neck region also acts as a pH-sensor in physiological conditions to regulate DC function. (Tabarani, Thepaut et al. 2009) The extracellular C-terminal domain is required for glycoprotein binding. DC-SIGN plays a critical role on almost every aspect of normal DC function, including antigen recognition and binding, antigen endocytosis and presentation, DC migration, and DC-T cell interaction. (Geijtenbeek, Krooshoop et al. 2000; Engering, Geijtenbeek et al. 2002; Frison, Taylor et al. 2003; Guo, Feinberg et al. 2004; Feinberg, Guo et al. 2005)

2.2 Specific Aims

Aim 1: To dissect the genetic architecture of the DC-SIGN gene in the Tobago population.

Aim 2: To test whether DC-SIGN genetic polymorphism is associated with elevated HHV-8 seroprevalence in the Tobago population.

Aim 3: To test whether DC-SIGN genetic polymorphism is associated with elevated risk of PC in the Tobago population.

2.3 Experiment Design

We chose a sample of 41 controls and 39 PC cases, and carried out a comprehensive genotypic screening, covering the entire DC-SIGN gene region (5' UTR, 7 exons, 6 introns and 3'UTR) using DHPLC and ABI BigDye sequencing techniques. Using the data we collected, we examined linkage disequilibrium (LD) and haplotype structure of DC-SIGN the Tobago population to accomplish Aim 1.

Based on the polymorphism and LD information from the initial screening, we further chose the most informative SNPs and potentially important regions, and extensively genotyped these SNPs and regions in a larger group of case and control samples. We achieved Aim 2 and 3 by case-control analysis based on the genotypic data we obtained.

2.4 Materials and Methods

2.4.1 Subjects

Subjects were drawn from the Tobago Prostate Survey, an ongoing population based screening study of men 40-79 years old on the Caribbean Island of Tobago. (Bunker et al. 2002) Recruitment was by word of mouth, posters, fliers, public health announcements and healthcare workers. A total of 2582 men aged 40 to 79 years old (50% of all Tobago men in this age group) were screened using both serum prostate-specific antigen (PSA) and digital rectal examination (DRE). Prostate cancer cases are further ascertained in men with elevated PSA level (4ng/mL or greater) or abnormal DRE by ultrasound-guided sextant biopsy.

Prior to screening, all participants gave written informed consent following procedures approved by the University of Pittsburgh and the Tobago Division of Health and Social Services institutional review boards. Fifteen milliliters of fasting peripheral blood was drawn prior to a

physician administered digital rectal exam (DRE), and prostate specific antigen (PSA) measured by the University of Pittsburgh Medical Center Pathology Laboratories. All subjects with an elevated PSA level ($\geq$4 mg/ml) or abnormal DRE findings (except simple enlargement) were referred to the Tobago Regional Hospital for six core, ultrasound-guided biopsy by surgeons trained by urologists from the University of Pittsburgh. Details are given in Bunker et al (2002; 2004).

HHV-8 seropositivity was determined by either indirect immunofluorescence assay (IFA) or a combination of enzyme-linked immunosorbent assay and IFA by Dr. Frank Jenkins. Details are given in Jenkins et al (J Infect Dis. 2004, 189:15-20).

Initial screening of 80 individuals (41 controls and 39 cases) was followed by sequencing 2kb of the 5' region of DC-SIGN in additional 153 cases and 172 controls.

2.4.2 ABI Sequence Analysis of DC-SIGN

Genomic DNA was isolated from peripheral blood leucocytes by standard procedures (Puregene, Gentra Systems). To obtain sufficient DNA, samples were whole genome amplified using the method of Dean et al. using the Genomiphi (Amersham Biosciences) whole genome amplification protocol. (Dean, Hosono et al. 2002) Briefly, 1 µl of DNA, at a concentration of approximately 40 nanograms/µl, was added to 9 µl of sample buffer containing random hexamers and heated to 95°C to denature the DNA. The sample was cooled and mixed with 9 µl of reaction buffer containing salts and deoxynucleotides and 1 µl of Phi 29 DNA polymerase enzyme mix. The mixture was incubated overnight at 30°C. After amplification, the Phi29 DNA polymerase was heat-inactivated by a 10 minute incubation at 65°C. All DNA samples used in this project were provided by Dr. Bunker.

2 kb of the 5 -promoter region, the exons, and 1 kb of the 3 -UTR were screened. All exons, exon-intron boundaries, proximal 5' and 3' untranslated regions were thoroughly covered. Unique sequence primers for each region were designed using MACVECTOR software. (www.macvector.com)

Amplifications were carried out by using Taq DNA polymerase kit (Invitrogen Corporation, CA). 20 ng of DNA was added to a reaction mixture, including 4pmoles of each primer pair (Table 1), 10nmoles of each dNTP, 2µl of 10xbuffer, 1 µl of $MgCl_2$ (Table 1), and 1U of Taq DNA polymerase. The final total reaction volume was adjusted to 20 µl by adding deionized water. Samples were denatured at 95°C for 2 min. followed by cycles [table1] of denaturation at 95°C for 30 sec, annealing at temperatures indicated in Table 1 for 30 sec. and extension at 72°C for 1 min. All amplimers (5µl) were resolved by electrophoresis in 2% agarose gel, stained with ethidium bromide and visualized under UV illumination to confirm the size and purity of the products.

The ExoSAP process was used to clean up the PCR products before sequencing. Briefly, 5µ PCR product was incubated with 5µ master mix of ExoSAP (0.5µl shrimp alkaline phosphatase, 0.05µl of exonuclease I, 0.5µl buffer and 4ul deionized water) at 37°C for 30 min. followed by denaturation at 88°C for 15 min. Amplimers were sequenced from both directions using the ABI Bigdye Terminator Ready Reaction kits (Applied Biosystems, Forest City, CA) as described in the manufacturer's protocol. Briefly, 2.5µl PCR product, 2.3µl primer (0.8µM), 0.575ul dRhodamine dye and 1.725µl of sequencing buffer together were mixed together, and the reaction mixture was incubated at 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 4 minutes, repeated for 24 cycles. Products were cleaned by ethanol precipitation procedures. Briefly, 5µl of EDTA and 60µl of 100% ethanol was added to each well of the optical 96-well

plate. The mixture was incubated at room temperature for 15min and centrifuged at 2500xg for 30min at 4°C and supernatant was removed. The DNA then was washed with 70% ethanol and centrifuged at 1650xg for 15min at 4°C. The fragments were separated on an ABI-3630XL capillary sequencer. All sequence analysis was carried out in the University of Pittsburgh Genomics and Proteomics Core Laboratory (www.genetics.pitt.edu). Sequences were aligned and curated using the program SEQUENCHER (Gene Codes Inc).

2.4.3 DHPLC

The exon 4 repeat variant was analyzed by denaturing high performance liquid chromatography (DHPLC) using a Varian Helix System. (Varian Associates Inc.)

Denaturing High Performance Liquid Chromatography (DHPLC) (Underhill, Jin et al. 1997) is a relatively recent technique for the detection of single base substitutions and small insertions/deletions in PCR products. The DHPLC system (Varian, Inc., Walnut Creek CA; Helix System) uses Ion-Pair Reverse Phase (IP-RP) chromatography to separate out DNA fragments. In HPLC, a stationary phase contained in the column is used to retain samples, and a mobile phase is used to release the sample off the column. For DNA IP-RP applications, the stationary phase is comprised of hydrophobic media. The mobile phase is a mixture of an aqueous buffer with an organic co-solvent (acetonitrile) and a counter ion of opposite charge from the sample DNA molecule (triethylammonium acetate – TEAA). The TEAA counter ion in the buffer is positively charged and forms pairs with the negatively charged phosphate groups of DNA, thus coating the DNA molecule in a hydrophobic layer. The number of TEAA molecules attached to the DNA is proportional to the length of the DNA fragment, therefore determining the degree to which the DNA is retained by the stationary phase. The coated DNA molecule is

Table 1: Amplification primer sequences and sequencing reaction components and conditions for DC-SIGN

Amplified region	**Forward Reverse**	**Primers**	**Annealing Temp (°C)**	**Mg2+ (mM)**	**Cycles**
Prom 1	PR1F	TGGTTCCTTGGAGTCACTCATGTC	59	2	35
	PR1R	CATGCACTGTGGAGGCAGACTG			
Prom2	PR2F	CCTGCTGGTTTCTTCACATCATC	60	2.5	35
	PR2R	TTCCTGCTGACTTTGGGAGTTG			
Exon 1 and 2	1&2F	GCAGGAGTTCTGGACACTGG	60	2	35
	1&2R	AGGAACCCAGGCCCAAGT			
Exon 3	3F	GGTGCTGAGGGATTTAGACC	58	2	35
	3R	ACACTGACACCTGGAGAAGG			
Exon 4	4F	AGGAGGAGGAGGACAGGAAG	59	2	35
	4R	TCAGAACAGAAGGCAGGAGACTG			
Exon 5	5F	TGTGATCATTGCCCTCAGTG	60	2	35
	5R	TCCCTCATTCATATCCTTCTCC			
Exon 6	6F	GCTTCTGGCTCCAACACATT	60	2.5	35
	6R	GCTGTCTGGAAGTGGGTATGA			
Exon 7	7F	GAGAAGAGAGGCAGGAACTCAAGG	60	2.5	35
	7R	GACAAGAAGGACAGAATGGGACC			
3' UTR	3' F	TGGATCTGCAAAAAGTCCGC	59	2.5	35
	3' R	AGGAGAGGCACAATTCACCTAGC			

absorbed to the column matrix until it is exposed to a specific concentration of acetonitrile in the mobile phase. The application of an increasing acetonitrile gradient releases the sample in order of increasing length, with smaller fragments eluting first and larger fragments eluting later. The DNA that is eluted from the column is detected by ultra-violet light absorption at a wavelength of 260nm using a uv/vis detector supplied with the DHPLC apparatus.

2.4.4 TDI-FP Genotyping

TDI-FP stands for Template directed Dye Terminator Incorporation assay with detection by Fluorescence Polarization. It is a single base primer extension assay coupled with homogeneous FP detection. The TDI-FP protocol was first reported in 1999. (Chen, Levine et al. 1999) There are four major steps to this assay: (i) template amplification by PCR; (ii) PCR product clean-up by Exo-SAP-It reagent; (iii) single-base primer extension using a primer that anneals one base 5' of the polymorphic site and incorporates fluorophore labeled terminators; (iv) FP reading and data analysis. *(Fig 2)*

PCR amplifications were carried out in a total reaction volume of 20 µl, containing 20ng of DNA, 4 pmoles of each primer, 4 nmoles of each dNTP, 2.0ul of 10xbuffer, 0.8µl of 25mM Mg^{2+} (Table 2) and 0.9U Taq DNA polymerase. The amplification and amplicon check procedures were the same as described before. Excess PCR primers and unincorporated dNTPs were degraded by ExoSAP process (USB Corporation, Cleveland, Ohio, USA). 10µl PCR product was incubated with a 10µl master mix of ExoSAP (1µl shrimp alkaline phosphatase, 0.1µl of exonuclease 23I, 1µl buffer and 8µl deionized water) at 37°C for 90 min., followed by denaturation at 95°C for 15 min.

The TDI-FP assay previously described (Chen, Levine et al. 1999; Kwok 2002) was then performed for high throughput genotyping. The cleaned PCR product (total volume 20µl) was combined with 0.05 µl Thermosequenase DNA polymerase (USB corporation, OH), 1.0µl 10X reaction buffer, 0.05µl of the 1:16 two-dye mix (eg. for SNP C/T, 16µl 0.1mM ddATP, 16µl 0.1mM ddGTP, 15µl 0.1mM ddCTP, 16µl 0.1mM ddTTP, 1µl R110-UTP, 1µl TAMRA-CTP) (Perkin Elmer Life Sciences, Inc, Boston, MA), 1.0µl 10µmol/L SNP specific, and 8µl water. Template-directed incorporation (TDI) was performed in 96-well, black-skirted plates (MJ Research, Waltham, MA) using an Eppendorf cycler under the following conditions: denaturation at 95°C for 1 min, followed by 35 cycles of denaturation at 94°C for 10s and simultaneous annealing and extension at the appropriate temperature (table 2) for 30s with a final hold step of 4°C. FP was measured by L.J.L Biosystem's Analyst HT Assay Detection System and analyzed by Allele Caller software package (L.J.L Biosystems, Sunnyvale, CA). The amplification and TDI-FP primers and conditions were summarized in Table 2.

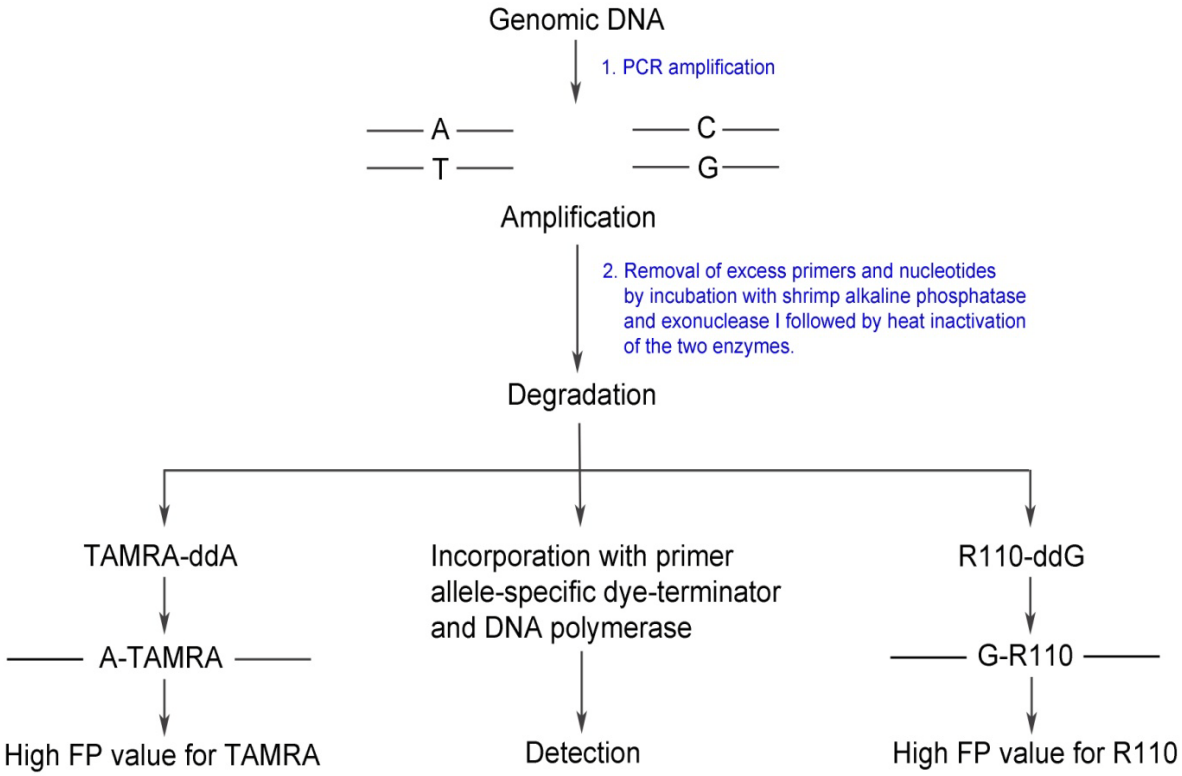

Figure 2: TDI-FP assay Principle and Procedures

Dideoxynucleotide dye-terminators were used in our assay.

Table 2: Amplification primer sequences and conditions and FP detection primer sequences for each FP locus

SNP ID (RS#)	**Position (relative to DC-SIGN start codon)**	**Primers**		**Annealing Temp(°C)**	**Mg2+ (mM)**
rs735240	-939	PCR-F	TCTCCACCTGCTCCTTCCTGATAA	60	2
		PCR-R	GTGTCCACCAGCAGGTGAATGATA		
		FP-F	CTGTGGTCCCCAGGAGTCCTG	55	
		FP-R	CCCCAAGACCCAAGAATCTGG		
rs4804806	-4228	PCR-	AAGACAGCCATCAGCACTACGG	60	2.5
		PCR-R	GCACTGACTGGTCTCCAACTTAGG		
		FP-F	GAACATTCGGCTGTAGTGCCATGT	55	
		FP-R	TGGTCAGAGTGGCTGGTCCCACAT		
rs1010047	18895	PCR-F	AAGGGCCTGTCTCAGGATGCCCATA	59	2
		PCR-R	TGGGACCGCTGCATTCTGGTCTA		
		FP-F	GGACACCTGAAGCATTTGGAAGGG	55	
		FP-R	TCCCTCTCCCCAAAGACAAAGGTG		

2.5 Data Analysis

2.5.1 Single SNP association test

For single SNP, we tested its genotype frequencies for fit to Hardy-Weinberg equilibrium expectation in cases and controls, using the χ^2 test with one degree of freedom.

2.5.2 Linkage Disequilibrium

The linkage disequilibrium (LD) relationships between the SNPs were investigated using the analysis program Ldmax, and the results displayed using the GOLD program. Ldmax uses the expectation-maximization algorithm of Slatkin and Excoffier (1995) to estimate linkage statistics. Haplotype frequencies for all marker pairs are calculated, and a table summarizing LD statistics in the region can be output, and displayed graphically by software GOLD. (Fig. 3) (http://www.sph.umich.edu/csg/abecasis/gold/index.html)

2.5.3 Haplotype construction and association analysis

Haplotype construction and case-control significance tests were done using a Bayesian MCMC algorithm, implemented in the PHASE software package. (Stephens, Smith et al. 2001) (http://stephenslab.uchicago.edu/software.html)

2.6 Results

2.6.1 DC-SIGN sequence variations

From the pilot screening, a total of 40 single nucleotide polymorphisms (SNPs) were identified, as shown in Table 3. There is only one SNP (within exon 7) in the coding sequence of DC-SIGN, and this does not cause an amino acid change. Most variations were found in the

5’ UTR, and SNPs in 5’UTR were in LD with each other. (Fig 3) Variations in the regulatory elements of DC-SIGN may affect the transcriptional activity and thus contribute to the elevated risk of viral infection or PC. Therefore a detailed screening of 2kb of the promoter region of DC-SIGN was carried out in 153 PC cases and 172 controls from Tobago.

Table 3: Locations and allele frequencies for SNPs within a 6Kb region across DC-SIGN in 80 Tobago samples

Relative position[1]	**Location**	**Minor Allele Frequency (MAF)**	**Allele**	**Reference Sequence #**
-1880	5'UTR	26.90%	G/A	rs8111321
-1827	5'UTR	0.50%	C/T	This report[2]
-1706	5'UTR	9.90%	T/C	rs11465355
-1605	5'UTR	1.10%	T/C	rs11465356
-1530	5'UTR	2.70%	A/C	rs11465352
-1509	5'UTR	2.20%	G/A	rs7392155
-1466	5'UTR	4.40%	C/T	rs4804804
-1275	5'UTR	1.10%	G/A	rs1146772
-1180	5'UTR	27.40%	A/T	rs7359874
-1089	5'UTR	0.50%	C/A	rs11465360
-1026	5'UTR	1.60%	C/T	This report[2]
-939	5'UTR	28.30%	C/T	rs735240
-336	5'UTR	49.40%	T/C	rs4804803
-201	5'UTR	6.60%	C/A	rs11465366

Table 3 continued.

Relative position[1]	Location	Minor Allele Frequency (MAF)	Allele	Reference Sequence #
-139	5'UTR	18.20%	C/T	rs2287886
216	Intron 1	44.10%	C/G	rs7252229
262	Intron 1	7.60%	G/T	rs11469369
451	Intron 1	2.90%	G/T	rs11465371
467	Intron 1	7.10%	G/A	rs11465322
3070	Intron 5	19.90%	G/A	rs8105572
3114	Intron 5	14.00%	G/C	rs8105483
3173	Intron 5	0.50%	C/T	This report[2]
4234	Intron 6	38.80%	C/G	rs11465391
4238	Intron 6	6.30%	G/T	rs11465392
4442	Exon7	9.30%	C/T	rs11465393
4655	3'UTR	1.10%	C/G	rs11465395
4787	3'UTR	22.10%	C/T	rs4084802
4847	3'UTR	13.60%	C/T	rs10403018
4915	3'UTR	34.30%	A/T	rs480410
4920	3'UTR	0.50%	A/T	This report[3]
4939	3'UTR	1.60%	T/C	rs11465396
5147	3'UTR	18.90%	C/T	rs11465397
5194	3'UTR	31.90%	G/A	rs6603119
5226	3'UTR	33.00%	C/T	rs7248772

Table 3 continued.

Relative position[1]	Location	Minor Allele Frequency (MAF)	Allele	Reference Sequence #
5238	3'UTR	3.30%	G/A	rs11465400
5282	3'UTR	12.20%	C/T	rs11465401
5370	3"UTR	42.00%	C/T	rs7248637
5470	3'UTR	7.10%	C/T	rs11465402

[1] Numbered with reference to GENBANK entry NM021159.2

[2] These SNPs were not found from dbSNP, and no known SNPs nearby

[3] rs4804801 is registered 5bp from this locus, could represent the same SNP due to mis-labeling of its position.

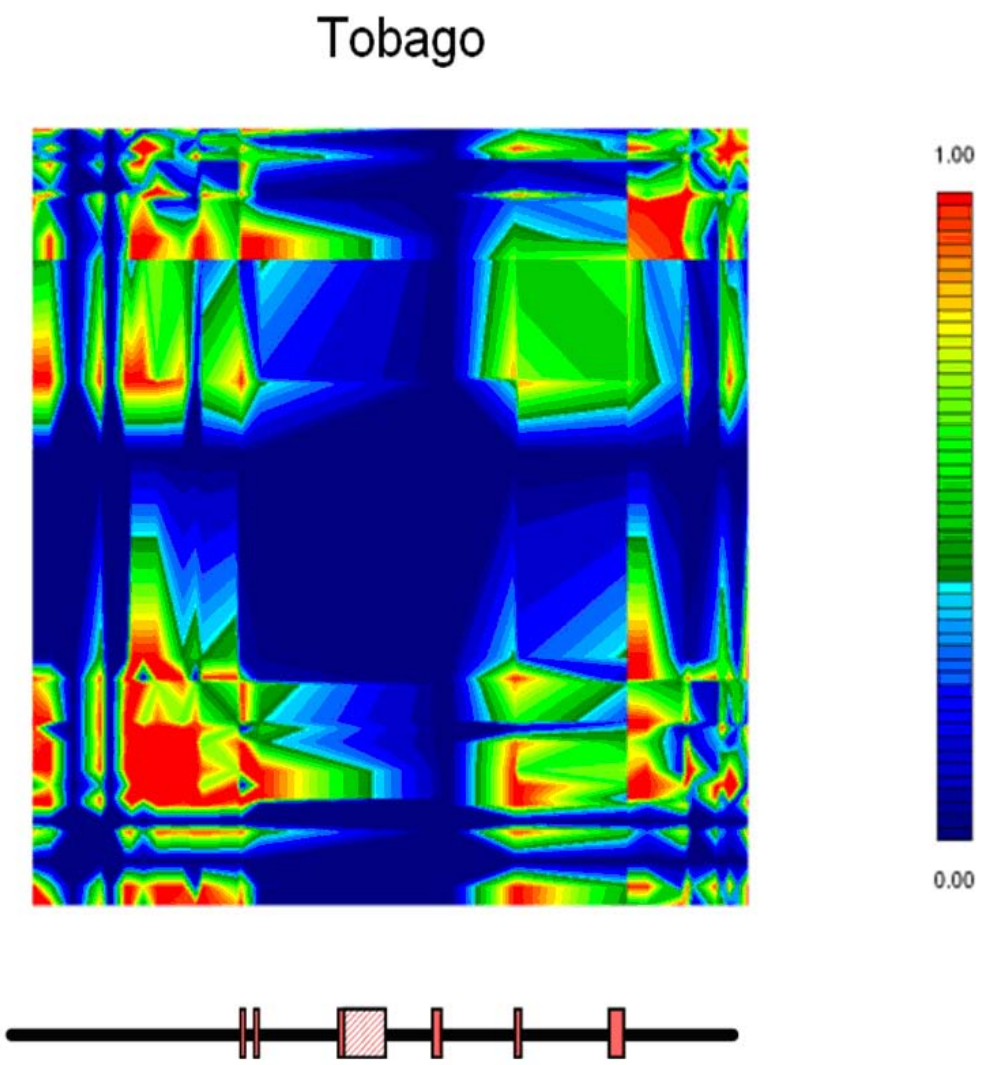

Figure 3: DC-SIGN SNP Pairwise D'LD plots in Tobago. Linkage disequilibrium plot for all SNPs from Table 3 is generated by GOLD software. LD on exon 4 repeat region is not included because no genotype data is available. High LD regions (red blocks) are mostly clustered in 5' and 3' UTR of the gene.

2.6.2 DC-SIGN Promoter region variations

A total of 26 promoter region SNPs were identified in our study, of which 17 have a minor allele frequency (MAF) of greater than 1%. (Table 4) The SNPs are in high LD with each other. (Fig 4) A total of 5 SNPs have a minor allele frequency over 10%. We ran both allele and genotype association tests on these 5 common SNPs, and none gave a statistically significant result at the cut-off p-value of 0.05, although two gave a marginally significant p value. (Table 5 and Table 6) This suggested that combination of SNPs among multiple loci, rather than a single SNP, might be important. Therefore we used the 5 Tag SNPs from Table 5 to generate promoter haplotypes of DC-SIGN. An association analysis was performed for haplotype distributions among cases and controls using PHASE software.

Table 4: Locations and allele frequencies for SNPs within the 2kb promoter of DC-SIGN in 325 Tobago samples

Position	Minor Allele freq	Allele	Reference Sequence number
-1880	24.50%	G/A	rs8111321
-1706	6.10%	T/C	rs11465355
-1605	1.00%	T/C	rs11465356
-1530	1.70%	A/C	rs11465357
-1509	1.40%	G/A	rs73921551
-1466	1.80%	C/T	rs4804804
-1463	1.50%	G/A	N/A
-1275	0.80%	G/A	rs11465359

Table 4 continued.

Position	Minor Allele freq	Allele	Reference Sequence number
-1180	29.10%	A/T	rs7359874
-1089	0.10%	C/A	rs11465360
-1026	2.10%	C/T	N/A
-939	28.30%	C/T	rs735240
-899	0.40%	C/T	N/A
-871	1.70%	T/C	rs735239
-866	0.10%	T/C	N/A
-819	1.70%	T/C	rs8112852
-802	0.80%	A/T	N/A
-745	6.10%	C/A	rs11465362
-670	0.30%	G/A	rs11465363
-601	5.80%	C/T	rs11465364
-462	0.30%	T/A	N/A
-460	0.60%	C/T	rs11465365
-336	44.10%	T/C	rs4804803
-201	7.60%	C/A	rs11465366
-171	0.30%	G/A	N/A
-139	24.50%	C/T	rs2287886

Table 5: Tag SNPs used to construct DC-SIGN promoter haplotypes

Position	Rs#	Alleles	Frequency in cases (2N=344)	Frequency In controls (2N=306)	p-value for each SNP frequency among cases and controls	OR (95% CI)
-1880	rs8111321	G	265(0.770)	227(0.741)	0.45	1.31 (0.59–2.89)
		A	79(0.230)	79(0.259)		
-1180	rs7359874	A	250(0.727)	212(0.693)	0.23	1.38 (0.70–2.73)
		T	94(0.273)	94(0.307)		
-939	rs735240	C	260(0.756)	208(0.681)	0.17	1.83 (0.87–3.87)
		T	84(0.244)	98(0.319)		
-336	rs4804803	T	176(0.513)	173(0.565)	0.11	0.60 (0.33–1.07
		C	168(0.487)	133(0.435)		
-139	rs2287886	T	84(0.245)	75(0.244)	0.83	0.98 (0.44–2.18)
		C	260(0.755)	231(0.756)		

Abbreviations: CI = confidence interval; OR = odds ratio.

Numbers of chromosomes (2N analyzed).

Table 6: Genotype association with PC for the 5 common SNP loci, which were used to generate haplotypes

Position	**Genotype**	**Cases N (%)**	**Controls N (%)**	**OR (95% CI)**
-1880	G/G	93 (60.8%)	97 (56.4%)	1.00 (Ref)
	G/A	49 (32.0%)	60 (34.9%)	1.17 (0.73-1.88)
	A/G	11 (7.2%)	15 (8.7%)	1.31 (0.57-2.99)
-1180	A/A	85 (55.6%)	88 (51.2%)	1.00 (Ref)
	A/T	53 (34.6%)	62 (36.0%)	1.13 (0.70-1.81)
	T/T	15 (9.8%)	22 (12.8%)	1.42 (0.69-2.91)
-939	C/C	90 (58.5%)	82 (47.4%)	1.00 (Ref)
	C/T	52 (34.0%)	71 (41.3%)	1.49 (0.94-2.39)
	T/T	11 (7.2%)	19 (11%)	1.89 (0.85-4.22)
-336	T/T	50 (32.7%)	62 (36.0%)	1.00 (Ref)
	T/C	58 (37.9%)	78 (45.3%)	1.08 (0.66-1.80)
	C/C	45 (29.4%)	32 (18.6%)	0.57 (0.31-1.03)
-139	T/T	13 (8.5%)	15 (8.7%)	1.00 (Ref)
	T/C	49 (32.0%)	55 (32.0%)	0.97 (0.42-2.24)
	C/C	91 (59.5%)	102 (59.3%)	0.97 (0.44-2.15)

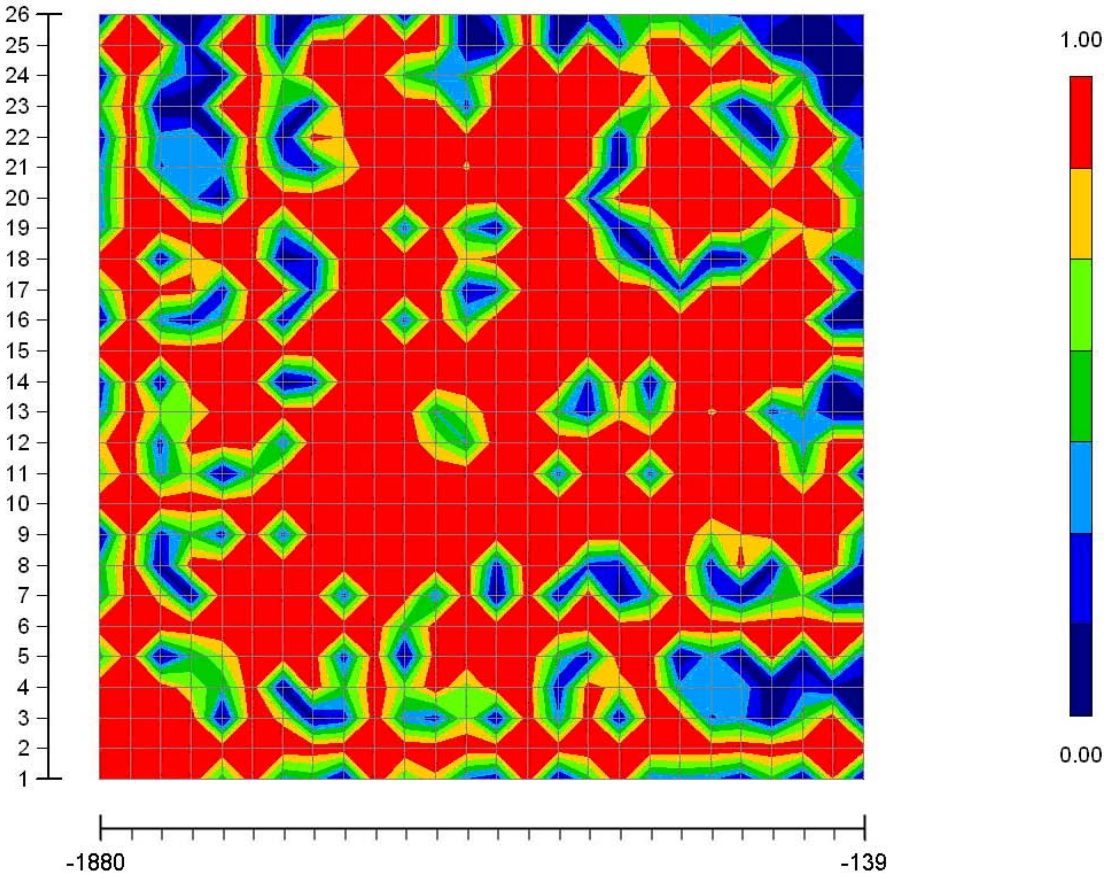

Figure 4: Linkage disequilibrium plot for the 26 SNPs within DC-SIGN promoter region
Linkage disequilibrium plot for all SNPs from Table 4 is generated by GOLD software. Red block indicates high LD, thus this figure shows high LD throughout the entire promoter region.

2.6.3 DC-SIGN Promoter Haplotype Association with PC

The haplotype distributions between PC cases and controls were significantly different with a p-value of 0.03. (Table 7 and Figure 5) This was mainly due to highly divergent frequencies of two haplotypes GATTC and GACTC that differ only by a single C/T SNP rs735240. (Table 7 and Figure 5) However, none of the odds ratio tests for the 5 individual SNPs were statistically significant at cut-off p-value of 0.05. (Table 5) Haplotype "GATTC" and "GACTC" showed significant association with PC as shown in Table 7. However, we calculated p-value

for each haplotype assuming all the haplotype frequencies are observations instead of estimates, therefore we need to be more conservative on those association results. These observations, combined with the fact that distributions of other haplotypes containing rs735240 are not distinct between cases and controls, suggest that complex haplotype structures rather than a single SNP may play an important role in PC susceptibility.

2.6.4 DC-SIGN Promoter Haplotype Association with HHV-8

While this DC-SIGN promoter haplotype is statistically significantly associated with elevated risk of PC in Tobago population (p=0.03), we did not find any haplotype association with HHV-8 serological status (p=0.28). We also did a stratified analysis for haplotype distribution among PC cases based on their HHV-8 serological status using PHASE software, and no significant difference was detected for haplotype distribution among PC cases with and without HHV-8, suggesting that no interactions occurred. (data not shown) These results suggest that association of DC-SIGN promoter haplotypes with prostate cancer risk is not due to their role in HHV-8 infection. Other pathogens, or other genes in LD with DC-SIGN promoter haplotype, may be causal to such association.

2.6.5 Extended genotypic association

To test whether the association we found was due to LD with other nearby causal loci, two additional tag SNPs from HapMap were picked up to extend the coverage into nearby chromosome regions. These were genotyped in cases and controls in order to test whether the haplotype association is due to LD with nearby genes or loci. SNP rs1010047 is located near the LSEctin gene, while SNP rs4804806 is located in the gap between DC-SIGN and DC-SIGNR.

Both these SNPs were in LD with rs735240 (-939), which is the major locus that contributes most to the haplotype association we found with PC in Tobago. (Figure 6) Both SNPs were genotyped by FP and the allelic frequencies were summarized in Table 8.

Table 7: DC-SIGN haplotype frequencies in prostate cancer cases and controls

Haplotype	**Freq in controls**	**Freq in cases**	**p-value[2]**
GACCC	0.3207	0.3885	0.25
GACTT	0.1825	0.1439	0.43
GATTC	0.1667	0.0259	0.01
ATTTC	0.1068	0.0563	0.14
ATCCC[1]	0.0738	0.014	N/A
GACTC	0.0523	0.1683	0.02
ATTTT[1]	0.0242	0.0475	N/A
ATCTT[1]	0.0223	0.0066	N/A
GTCCC[1]	0.0222	0.0299	N/A
ATTCC[1]	0.0062	0.0676	N/A
GATCC[1]	0	0.0128	N/A
GTTTT[1]	0	0.0133	N/A

[1]Haplotype with a frequency of <5% in cases and controls were pooled, and p-valued are not estimated for these subgroups.

[2] Null hypothesis: haplotype frequencies are the same between cases and controls

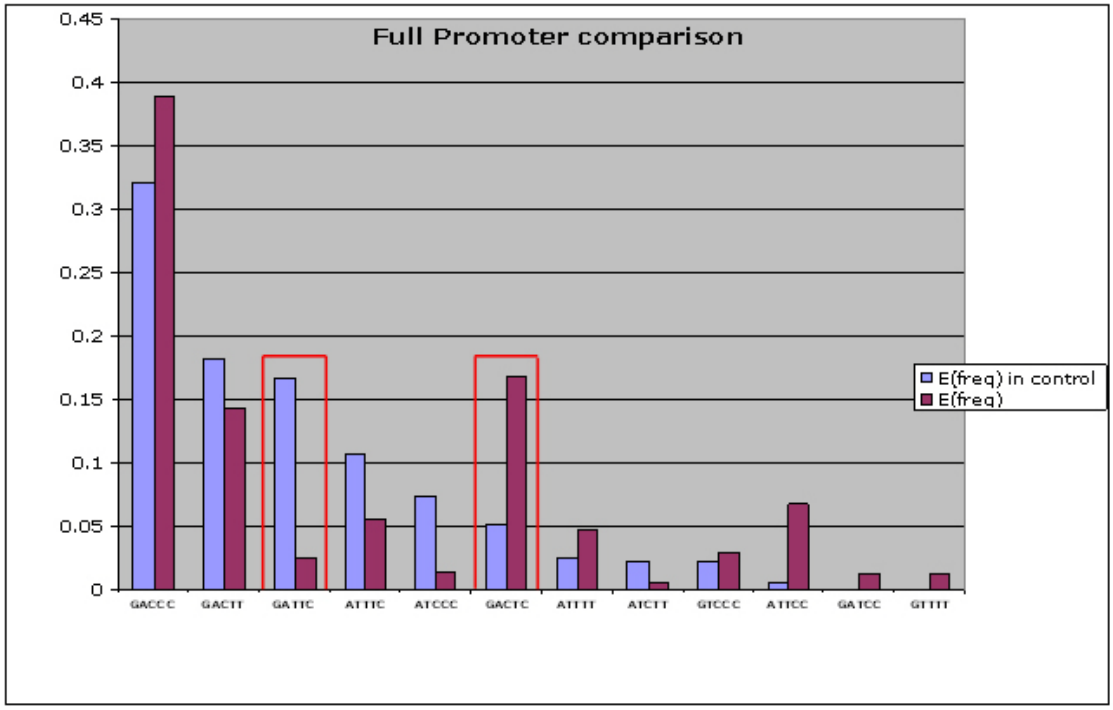

Figure 5: Haplotype distribution in prostate cancer cases and controls. Frequencies for each haplotype were sorted and displayed in the order of frequencies in controls. Two haplotypes (GATTC and GACTC), which differ most between cases and controls are highlighted in red.

Both SNPs are in high LD with rs735240 (-939) in our samples, which is consistent with the Hapmap result. Genotypic association with PC for each individual SNP was summarized in Table 9. While we did not see any association for rs1010047, we found a statistically significant association with PC for rs4804806. Odds ratio of genotype GG to GA genotype is 0.47 with 95% CI of (0.28-0.81), and odds ratio of genotype GG to AA is 0.34 with 95% CI of (0.18-0.64). This suggests that rs4804806 is significantly associated with elevated risk of PC, and that the A allele of this locus appears to be a dominant risk allele for PC in Tobago population.

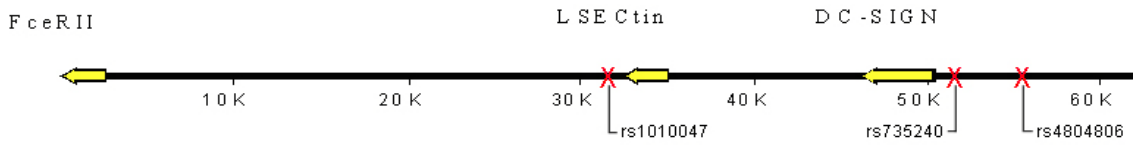

Figure 6: DC-SIGN and neighboring genes, and relative positions of the TAG SNPs we used for extended genotyping and association analysis.

Table 8: Summary of allele frequencies of the two extended SNPs

position	allele	MAF in all	MAF in controls	MAF in cases	reference sequence number
18895	A/G	38.9% (G)	40.40% (G)	38.80% (G)	rs1010047
-4228	G/A	46.6% (A)	40.10% (A)	53.90% (A)	rs4804806

Table 9: Genotype association with PC for 2 extended SNPs

Position	Genotype	Cases N (%)	Controls N (%)	OR (95% CI)
rs1010047 (+18895bp)	G/G	25 (16.3%)	38 (22.1%)	1.00 (Ref)
	G/A	63 (41.2%)	63 (36.6%)	0.66 (0.36-1.22)
	A/A	65 (42.5%)	71 (41.3%)	0.72 (0.39-1.32)
rs4804806 (-4228bp)	G/G	31 (20.3%)	64 (37.2%)	1.00 (Ref)
	G/A	79 (51.6%)	78 (45.3%)	0.47 (0.28-0.81)
	A/A	43 (28.1%)	30 (17.4%)	0.34 (0.18-0.64)

2.6.6 Extended haplotype association with PC

rs4804806 is over 4kB upstream of the DC-SIGN gene, but it is still close enough to be within the 5' regulatory region of the gene. We therefore reconstructed the promoter haplotype adding rs4804806 in additional to the 5 SNPs previously used in Table 5. The extended haplotypes were summarized in Table 10, and the extended haplotype distribution among PC cases and controls is still significantly different (p-value=0.01). Such association is mostly caused by the added rs804806 SNP because most haplotypes with the A allele at this locus are overrepresented in cases than controls, as seen in Figure 7.

Table 10: DC-SIGN promoter region haplotype frequencies in prostate cancer cases and controls adding rs4804806

Haplotype	Freq in controls	Freq in cases
GGACCC	0.212	0.198
AGACCC	0.148	0.24
GGACTT	0.116	0.066
GGATTC	0.101	0.012
AGACTT	0.093	0.098
AGATTC	0.067	0.016
GATTTC	0.064	0.026
GATCCC	0.044	0
AATTTC	0.043	0.04
GGACTC	0.031	0.087
AATCCC	0.03	0.008

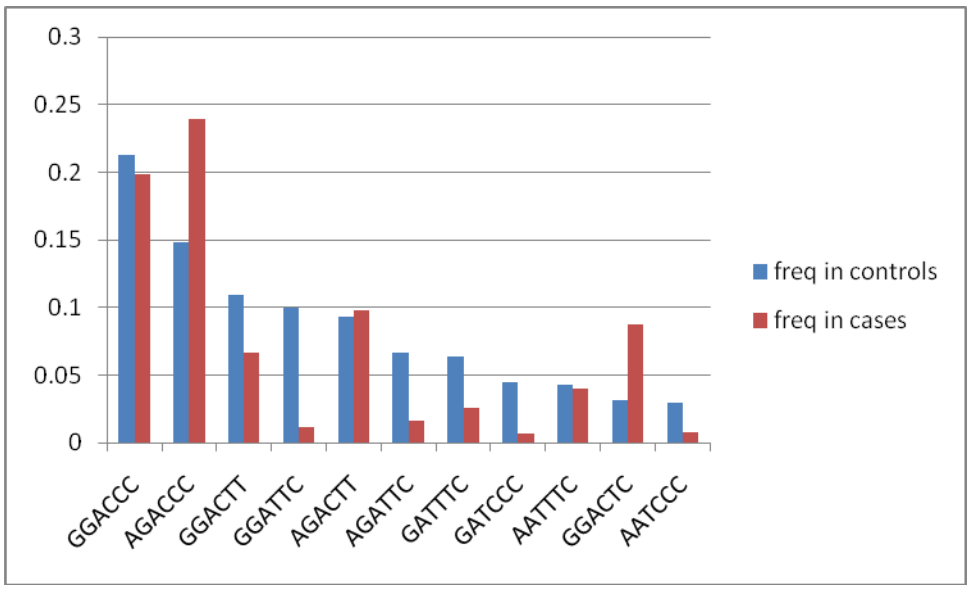

Figure 7: Haplotype distribution in prostate cancer cases and controls adding rs4804806 Frequencies for each haplotype were sorted and displayed in the order of frequencies in controls, only haplotypes with over 3% frequency in controls are listed. The most common haplotype in cases and controls (GGACCC in controls and AGACCC in cases) differed by rs4804806 G/A only. Most distortions in the haplotype distribution between cases and controls are caused by rs4804806 locus.

refSNP rs4804806 with alleles A/G in dbSNP b126 (dbSNP report | Ensembl SNPview)
chr19:7722625..7722625, (+) strand relative to the human reference sequence

Population	Genotype frequencies										Allele frequencies						
											Ref-allele			Other-allele			
	genotype	freq	count	genotype	freq	count	genotype	freq	count	Total	allele	freq	count	allele	freq	count	Total
ASW (A)	G/G	0.472	25	A/G	0.434	23	A/A	0.094	5	53	G	0.689	73	A	0.311	33	106
CEU (C)	G/G	0.301	34	A/G	0.584	66	A/A	0.115	13	113	G	0.593	134	A	0.407	92	226
CHB (H)	G/G	0.762	64	A/G	0.238	20	A/A	0	0	84	G	0.881	148	A	0.119	20	168
CHD (D)	G/G	0.729	62	A/G	0.247	21	A/A	0.024	2	85	G	0.853	145	A	0.147	25	170
GIH (G)	G/G	0.682	60	A/G	0.261	23	A/A	0.057	5	88	G	0.812	143	A	0.188	33	176
JPT (J)	G/G	0.698	60	A/G	0.267	23	A/A	0.035	3	86	G	0.831	143	A	0.169	29	172
LWK (L)	G/G	0.300	27	A/G	0.556	50	A/A	0.144	13	90	G	0.578	104	A	0.422	76	180
MEX (M)	G/G	0.300	15	A/G	0.520	26	A/A	0.180	9	50	G	0.560	56	A	0.440	44	100
MKK (K)	G/G	0.238	34	A/G	0.573	82	A/A	0.189	27	143	G	0.524	150	A	0.476	136	286
TSI (T)	G/G	0.307	27	A/G	0.523	46	A/A	0.170	15	88	G	0.568	100	A	0.432	76	176
YRI (Y)	G/G	0.336	38	A/G	0.531	60	A/A	0.133	15	113	G	0.602	136	A	0.398	90	226

Note: the 'reference' allele is the base observed in the reference genome sequence at this location

Population descriptors:
ASW (A): African ancestry in Southwest USA
CEU (C): Utah residents with Northern and Western European ancestry from the CEPH collection
CHB (H): Han Chinese in Beijing, China
CHD (D): Chinese in Metropolitan Denver, Colorado
GIH (G): Gujarati Indians in Houston, Texas
JPT (J): Japanese in Tokyo, Japan
LWK (L): Luhya in Webuye, Kenya
MEX (M): Mexican ancestry in Los Angeles, California
MKK (K): Maasai in Kinyawa, Kenya
TSI (T): Toscans in Italy
YRI (Y): Yoruba in Ibadan, Nigeria

Figure 8: rs4804806 genotype frequencies in different populations from Hapmap (http://hapmap.ncbi.nlm.nih.gov/cgi-perl/snp_details_phase3?name=rs4804806&source=hapmap27_B36&tmpl=snp_details_phase3)

2.7 Discussions

In this study, we show that there are few genetic variants that cause amino acid changes in DC-SIGN. This suggests that DC-SIGN is highly conservative through evolution, and this is consistent with other studies. (Barreiro, Patin et al. 2005) We found fewer SNPs in the coding region of DC-SIGN compared to Barreiro's group's finding in 2005 in sub-Saharan African

samples from HapMap, and the LD pattern was also slightly different. While this might be explained by different sample sizes, it also suggests that genetic characteristic of DC-SIGN in Tobago population may be different from that of the West African reference population. This may reflect the genetic differentiation present in African populations or founder effects due to the restricted sample size of the original population. Many SNPs are clustered in the 5'UTR of DC-SIGN and are in high linkage disequilibrium: this suggests that transcriptional modulation, rather than translational modulation may play an important role to alter DC-SIGN function.

The association of DC-SIGN promoter haplotypes with PC suggests a role for the innate immune system in the etiology of this disease. However, we don't see any significant association of DC-SIGN haplotypes with HHV-8 infection. It suggests that other pathogens may play a role in the elevated risk of PC in Tobago. In 2006, Urisman et at. identified a novel gamma-retrovirus, xenotropic murine leukemia virus related virus (XMRV) in prostate tumors of patients homozygous for R462Q RNASEL variant.(Urisman, Molinaro et al. 2006) This virus has been shown to be a PC risk factor and susceptibility mediated by the IFN antiviral pathway. (Dong, Kim et al. 2007) DC-SIGN regulates several retroviral infections. Therefore, the association we found between DC-SIGN promoter haplotypes and PC might be due to its regulatory role in determining susceptibility to XMRV infection in prostate tissues in the Tobago population. Further, our results cannot rule out the role of HHV-8 infection on elevated PC risk in Tobago as suggested by epidemiological studies. (Hoffman, Bunker et al. 2004) In 2008, Montgomery, Jenkins et al. reported that a SNP in the IL-6 gp130 signaling receptor is associated with increased PC risk among HHV-8 seropositive men in Tobago. (Montgomery 2008) This is suggests that HHV-8 may exploit genetic variations in genes other than DC-SIGN and contribute to PC risk in Tobago.

The DC-SIGN promoter SNP rs4804803 (-336 C/T) has been reported to be associated with viral infection. (Martin, Lederman et al. 2004; Despres, Sakuntabhai et al. 2005; Sakuntabhai, Turbpaiboon et al. 2005) The C allele of this promoter SNP showed increased SP1 binding activity in vitro and has been suggested to have a significant protective effect against severity of Dengue fever. (Sakuntabhai, Turbpaiboon et al. 2005) On the other hand the C-336 allele has been associated with elevated risk of parenteral acquisition of HIV-1 infection. (Martin, Lederman et al. 2004) In our study, we did not see any statistically significant association with this SNP. We only found marginal evidence for association at the C-336 and -939C/T loci. Although the functions of these two SNPs can be further analyzed by reporter gene assay, we find a much stronger association between haplotype distribution and PC risk, suggesting that the complex haplotype structure rather than any single SNP may be responsible for the elevated PC risk in the population of Tobago. The observed association may also be caused by linkage disequilibrium with other variation found in this genomic region.

The SNP rs4804806 (-4228) is the only locus which by itself showed significant association with PC risk in our study, and the A allele appears to be a risk allele for PC. The A allele frequency is 40.1% in our Tobago controls, which is quite similar to that seen in the African YRI samples from Hapmap (Fig 8). Hapmap data also showed that African populations have the highest frequencies for the A allele in all major populations, followed by Caucasians, then Asians. (Fig 8) Interestingly, risk of PC follows the same order among these ethnic groups. We searched for consensus sequence transcription factor binding motifs near this locus using TRANSFAC (www.gene-regulation.com), and found that the A allele will actually add a putative transcription factor GR-beta binding site. Therefore the association we found at this locus might be explained by its regulation of DC-SIGN transcription. However this needs to be

further explored since this locus is quite far away (over 4kb) from the gene, and there are several redundant putative GR-beta binding sites within the region 4kb upstream of DC-SIGN. Thus it is also possible that this is not the causal locus, and the haplotype association observed at this locus could be due to its linkage disequilibrium with other causal variation.

2.8 Conclusion

To conclude, in this study we found significant association between promoter haplotypes in the DC-SIGN region and elevated risk of PC in Tobago population, while we did not see any association with serological evidence of HHV8 infection. We also found significant association with PC for rs4804806 (G-4228A), although while its function still remains to be explored. Infectious agents other than HHV8, and other genes linked to DC-SIGN promoter haplotype, may be causal to elevated risk of PC in Tobago.

3.0 CHAPTER THREE: ASSOCIATION STUDY OF PSA RS266882 POLYMORPHISM AND ANDROGEN RECEPTOR $(CAG)_n$ REPEAT LENGTH POLYMORPHISM WITH PC RISK AND PSA LEVEL IN TOBAGO POPULATION

3.1 Background and Introduction

3.1.1 Androgen and androgen receptor

Androgens are essential for the growth and maintenance of both normal and malignant prostate cells. PC is believed to be androgen dependent, at least in early stages of the disease. Men who are castrated at an early age, or who have mutations that impair androgen production, do not develop prostate cancer. (1991; Wu and Gu 1991) It has been reported that inhibition of androgen signaling in vitro results in apoptosis, and androgen ablation therapy results in temporary tumor regression in 70-80% of PC cases clinically.(Eder, Culig et al. 2000) These studies suggest that the androgen and genes involved in androgen metabolism may play an important role in PC development.

The androgen receptor (AR) is a nuclear hormone receptor, and is the central molecule of the androgen signaling pathway. Ligand-dependent activation of AR by dihydrotestosterone and testosterone regulates a transcriptional cascade of many androgen dependent genes, including E-cadherin, the prostate-specific antigen (PSA) gene, APS gene, etc. The AR gene is located at

chromosome Xq11-12 and encodes a protein with two polymorphic regions in the transactivation domain of exon 1. Polyglycine repeats $(GGC)_n$ vary in length from 8 to 17, while polyglutamine repeats $(CAG)_n$ are more variable, with repeat lengths from 9 to 29 in Caucasians. The length of the CAG allele has been reported to be inversely related to transcriptional activity of AR, and several studies have suggested that short CAG alleles increase the relative risk of PC. (Giovannucci, Stampfer et al. 1997; Tut, Ghadessy et al. 1997; Xue, Irvine et al. 2000)

3.1.2 Prostate-specific antigen (PSA)

PSA is a serine protease produced by the secretory epithelial cells of the prostate gland. It can cleave insulin-like growth factor binding protein-3 and other proteins. PSA has been widely used as a diagnostic marker of PC since the early 1990s. Production of PSA is mediated by AR binding to the androgen response elements (ARE) in the promoter region of the PSA gene, located at chromosome 19q13.3-q13.4. Its promoter region contains several AREs to which AR can bind. A G/A polymorphism (rs266882) at -158bp, within ARE1 of PSA gene has been suggested to be associated with increased risk of PC and serum PSA level. (Xue, Coetzee et al. 2001; Cramer, Chang et al. 2003) Some studies further suggested that short CAG alleles of AR gene interact with the rs266882 G/G genotype, and such combination could significantly increase the relative risk of PC. (Xue, Irvine et al. 2000)

3.2 Specific Aims

To test whether the rs266882 G/A polymorphism in the AREI of the PSA gene, or the (CAG)n repeat polymorphism in AR gene, is associated with risk of prostate cancer or serum PSA level in men from the Tobago population, and to determine any evidence of gene x gene interaction between these two polymorphisms.

3.3 Experiment Design

This specific aim was accomplished through the following steps:

-Genotyping rs266882 G/A polymorphism in 167 cases and 320 controls drawn from the Tobago population, by TDI-FP.

-Genotyping the AR (CAG)n repeat polymorphism in the same samples.

-Calculating odds ratios (ORs) and confidence intervals (95% CI) to test if there is any association between PSA A/G genotype or AR (CAG)n repeat length and risk of PC, and to test whether there is any evidence of interaction between these two genes in determining the risk of PC in the Tobago population.

-Using linear regression approach to test if genotypes at either of these loci are associated with serum PSA levels in controls and to test for interaction between these two genes in determining PSA levels.

3.4 Subjects and Methods

3.4.1 Subjects

Subjects were drawn from the Tobago Prostate Survey, an ongoing population based screening study of men 40-79 years old on the Caribbean Island of Tobago. (Bunker et al. 2002) as previously described. Total of 167 PC cases and 320 age matched controls were used for this analysis.

3.4.2 TDI-FP Genotyping on Rs266882

PCR amplifications were carried out in a total reaction volume of 20 µl, containing 20ng of DNA, 4 pmoles of each primer, 4 nmoles of each dNTP, 2.0µl of 10xbuffer, 0.8µl of Mg^{2+} (Table 11) and 0.9U Taq DNA polymerase. The amplification and amplicon check procedures were the same as described before. Excess PCR primers and unincorporated dNTPs were degraded by the ExoSAP procedure (USB Corporation, Cleveland, Ohio, USA). 10µl PCR product was incubated with 10µl master mix of ExoSAP (1µl shrimp alkaline phosphatase, 0.1µl of exonuclease 23I, 1µl buffer and 8µl deionized water) at 37°C for 90 min., followed by denaturation at 95°C for 15 min.

A template directed dye-terminator incorporation with fluorescence-polarization (TDI-FP) assay, as described in the previous chapter, was then performed using the described protocol. The FP primers and conditions are listed in Table 11.

Table 11: Amplification primer sequences and conditions and FP detection primer sequences for PSA AREI SNP rs266882

SNPs ID (RS#)	Primers		Annealing Temp (°C)	Mg2+ (mM)
rs266882	PCR-F	TTGTATGAAGAATCGGGGATCGT	60	2
	PCR-R	TCCCCCAGGAGCCCTATAAAA		
	FP-F	ACTTGCTGTTCTGCAAT	55	
	FP-R	CCTTCCCCTCCCTCTCGATC		

3.4.3 AR $(CAG)_n$ repeat Genotyping

The AR (CAG)n was typed by polymerase chain reaction amplification from genomic DNA in the presence of a VIC-labeled forward primer, ARF: 5'-TCCAGAATCTGTTCCAGAGCGTGC-3'. Fragments were resolved on an ABI 3730 DNA analyzer (Applied Biosystems, Foster City, CA) and analyzed using GenMapper software version 3.7 (ABI). Fragment sizes were assigned by comparison to a sequence verified fragment ladder generated by combining the PCR products of individuals of known (CAG)n length. CAG repeat number ranged between 11 and 30 repeats in this sample (mean=19.7, SD=3.1), which is consistent with previously-reported CAG repeat lengths from populations of African ancestry. (Edwards, Hammond et al. 1992) The distribution in Tobago men was bimodal, with a nadir between 21 and 22 repeats. They were grouped for analysis into short (21 repeats or fewer) and long (22 repeats or more) subgroups.

3.4.4 Measurement of PSA level

Serum PSA levels were measured at the University of Pittsburgh Central Pathology Laboratory using the automated Microparticle Enzyme Immunoassay, Abbot AxSYM PSA assay (Abbott Laboratories, Abbott Park, IL). (www.abbottdiagnostics.com/Products/Instruments_by_Platform/default.cfm?system=AxSYM / AxSYM Plus&reg=us)

3.5 Results

3.5.1 Association analysis of PSA rs266882 genotype with PC

Genotype information for rs266882 and its association with PC are summarized in Table 12. We did not see any evidence for association at this locus with PC risk.

3.5.2 Association analysis of AR repeat length polymorphism with PC

The CAG repeat length follows the binomial distribution pattern in our samples. The mean CAG repeat length was 19.7 ± 3.1 (mean ± SD) in cases and 20.6 ± 2.8 in controls. Most studies used a cutoff of 22 repeats, to form two groups of repeat length ≤ 22 (short) and > 22 (long) repeats. The (CAG)n length averaged 20.6 in our control samples, and the repeat length follows a binomial distribution in our samples (Fig 9), and 20 is the nadir as shown in Figure 9. We therefore also used a cutoff of 21 in our analysis to make sure different cutoffs won't affect the results. As shown in Table 13, repeat length was not associated with PC risk in our sample, at either cutoff standard.

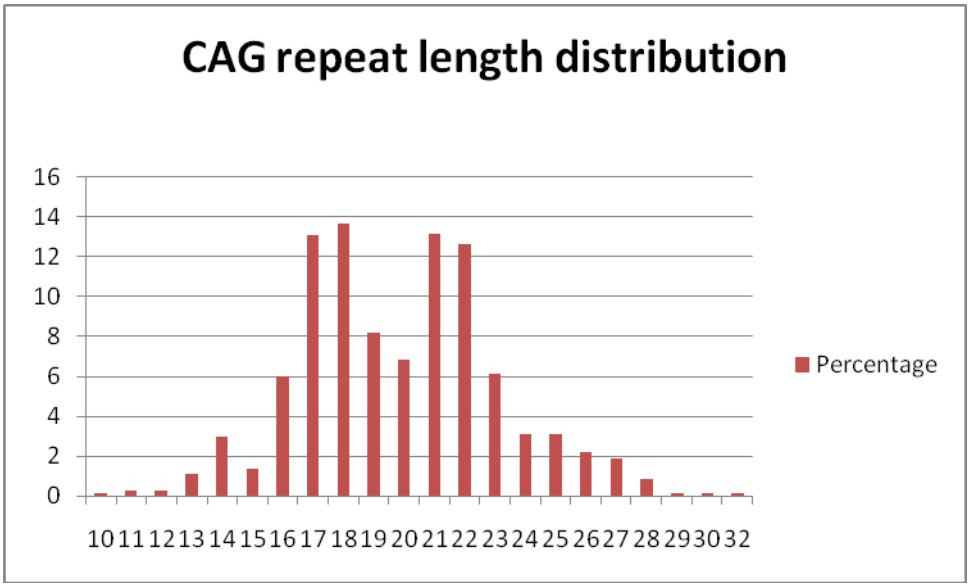

Figure 9: Histogram of CAG repeat length distribution

Table 12: ORs for PC associated with PSA rs266882

Locus	Genotype	Cases N (%)	Controls N (%)	OR[a] (95% CI)
PSA (rs266882, -158bp)	G/G	33 (19.8%)	72 (22.5%)	1.00 (Ref)
	G/A	79 (47.3%)	139 (43.4%)	0.81 (0.49-1.32)
	A/A	55 (32.9%)	109 (34.1%)	0.91 (0.53-1.53)

OR[a] adjusted for age

3.5.3 Association analysis of interaction between PSA rs266882 GG and AR (CAG)n repeat length polymorphisms in determining risk of PC

To test whether the GG genotype of PSA rs266882 interacts with short CAG repeat length and contributes to elevated PC risk together, as has been reported by some studies, we grouped our samples into two groups. One group was comprised of individuals with both the GG homozygote genotype and the short CAG allele, and the other group was comprised of all other

individuals. Our result suggested that there was no association with PC at cutoffs of 21 or 22. (Table 14)

Table 13: ORs for PC associated with AR CAG repeat length polymorphism

Locus	Genotype	Cases N (%)	Controls N (%)	OR[a] (95% CI)
AR (CAG)n (22 cut off)	Short	128 (76.6%)	225 (70.3%)	1.00 (Ref)
	Long	39 (23.4%)	95 29.7%)	1.38 (0.90-2.13)
AR (CAG)n (21 cut off)	Short	101 (60.5%)	186 (58.1%)	1.00 (Ref)
	Long	66 (39.5%)	134 (41.9%)	1.10 (0.75-1.61)

OR[a] adjusted for age

3.5.4 Regression analysis of PSA rs266882 genotype and AR CAG repeat length with serum PSA level

To test whether PSA rs266882 genotype or AR (CAG)n polymorphism is associated with serum PSA level in controls, we carried out a regression analysis in the 320 controls using a regression module implemented in R. We also added the interaction term in the formula to see whether these genetic polymorphisms interact with each other. Our results showed that neither PSA genotype nor AR (CAG)n repeat length is associated with serum PSA level in controls at the cut-off p-value of 0.05, and there is no interaction between these 2 loci. (Table 15)

Table 14: ORs for PC associated with AR CAG repeat length polymorphism

Locus	Genotype	Cases N (%)	Controls N (%)	OR[a] (95% CI)
AR (CAG)n (22 cut off)	GG+short	25 (15.0%)	53 (16.6%)	1.00 (Ref)
	Others	142 (85.0%	267 (83.4%)	0.89 (0.53-1.49)
AR (CAG)n (21 cut off)	GG+short	22 (13.2%)	47 (14.7%)	1.00 (Ref)
	Others	145 (86.8%)	273 (85.3%)	0.88 (0.51-1.52)

Table 15: Regression analysis for PSA level with PSA and AR genotypes

SL short=0, long=1, geno (aa=0, ag=1, gg=2)

Coefficients	Estimate	Std. Error	t value	p-value (>|t|)
pl$SL	-0.2286	0.15463	-1.478	0.14
pl$geno	-0.10501	0.08473	-1.239	0.216
pl$SL:p1$geno	0.12861	0.136	0.946	0.345

3.6 Discussion

Results from this study do not provide any evidence for an association with PC risk or PSA level for PSA -158 G/A genotypes or AR (CAG)n repeat length genotypes.

Given the importance of PSA, regulation of its production is of biological and clinical importance. Numerous studies have examined the function of rs26682 in PSA regulation, either alone or in combination with the AR gene. Some studies reported that the GG genotype is

associated with risk of PC (Xue, Coetzee et al. 2001; Cicek, Liu et al. 2005; Severi, Hayes et al. 2006), while some reported conversely that the A allele is associated (Gsur, Preyer et al. 2002; Medeiros, Morais et al. 2002; Lai, Kedda et al. 2007) The same genotype groups were included our study, and we failed to find any evidence of association, in agreement with other studies. (Wang, Sato et al. 2003; Salinas, Austin et al. 2005)

Although we did not find any association in this study, we cannot rule out the possible role of these phenotypes on PC risk in Tobago population. First, it is possible that the risk estimates may accurately reflect the small magnitude of the relationship between the AR or PSA polymorphisms and PC risk, but our sample size it not large enough to achieve statistical significance. Second, transactivation of genes by the AR is a complex process involving multiple other cofactors and other transcription factors, and influenced by the environment. Thus, there may be a true association, but only detectable in the presence of other gene-gene and gene-environment interactions.

3.7 Conclusion

No evidence was provided by our study that repeat length polymorphism in the AR gene or the rs266882 SNP in the PSA gene is associated with PC susceptibility or serum PSA level in the Tobago population.

4.0 CHAPTER FOUR: DISCUSSIONS

4.1 Summary

The two studies described here were designed to search for genetic variants that are associated with elevated risk of prostate cancer in the Tobago population.

The first study explored the relationship between an immune related candidate gene, DC-SIGN, and risk of prostate cancer. We initially genotyped the entire DC-SIGN coding and non-coding region in 41 controls and 39 cases from Tobago, and characterized the LD pattern of the DC-SIGN gene. Our pilot screening found that most informative SNPs were clustered in the promoter region of DC-SIGN, and were in high LD with each other. Therefore we further sequenced the 2kb promoter region of DC-SIGN in additional 153 cases and 172 controls from Tobago, and carried out association analysis for all the common SNPs we identified (MAF >10%) as well as determining their haplotype distribution. Although we only found marginal association for two individual SNPs (rs4804803 and rs735240), we did find significant evidence of association for promoter haplotypes of DC-SIGN and risk of PC (p-value=0.03). To test whether the haplotype association was due to LD with other nearby causal loci, two additional tag SNPs from HapMap were typed to extend the coverage into nearby chromosome regions. We found strong evidence of association for rs4804806 (-4228) itself (p-value < 0.01) as well as for promoter haplotype distributions with this locus. The A allele of this locus seems to be a dominant risk allele for elevated risk of PC, and the G allele seems to be protective against PC

risk. The function of rs4804806, and whether such association is due to linkage disequilibrium with other causal loci, remains to be explored. In the second study, we examined whether genetic variations of androgen receptor (AR) and PSA gene were associated with risk of prostate cancer or serum PSA level in the Tobago population. It has been reported that the rs266882 G/A polymorphism in the AREI of PSA gene, and the (CAG)n repeat polymorphism in androgen receptor (AR) gene, are associated with risk of PC and serum PSA level in several studies, but the results we obtained were equivocal.

In order to examine whether those two genetic variants affect PC risk or PSA level in the Tobago population, we genotyped rs266882 G/A, and CAG repeat length variants of AR, in 167 cases and 320 controls. Association analysis was conducted for each locus, and two genotypes were used to test for gene x gene interaction between these two loci. Regression analysis was carried out for each locus and combined with serum PSA level in 320 controls. Results from this study do not provide any evidence for association with PC risk or PSA level for PSA rs266882 G/A genotypes or AR CAG repeat length genotypes.

4.2 Future Directions

Our results suggested that the association for DC-SIGN promoter haplotype with PC cannot be explained by HHV infections. It is important to test whether DC-SIGN contributes to PC risk by regulating other pathogens. XMRV has been reported to be associated with PC risk. Assays to determine whether XMRV can use DC-SIGN as a receptor will be helpful. Alternatively, serological screening on other known DC-SIGN targeted infectious agents in our Tobago samples will also help to identify potential PC associated pathogens.

Further function analysis on the rs4804806 A/G polymorphism is also necessary. We propose reporter-gene assay in order to see if this SNP can alter promoter activity and regulate transcription. This variant affects a putative GR-beta transcription binding site. Therefore it is also helpful to test whether this variant has any effect on GB-beta binding in vitro.

We also propose to do association studies in African Americans and African populations with high PC risk. Such studies will provide us invaluable information on whether the DC-SIGN association we found in Tobago population is relevant in other African and African-decent populations.

4.3 Public Health Significance

Prostate Cancer is the most common cancer among men and the leading cause of cancer related deaths. It remains a heavy public health burden. We have shown a significant association of DC-SIGN promoter haplotypes with elevated risk of prostate cancer in the Tobago population. Our data strongly supports the hypothesis that chronic inflammation and genes involved in immune response play an important role in PC. The results from this dissertation lend support to exploring genetic causes and prevention of prostate cancer.

BIBLIOGRAPHY

"Urological oncology: reconstructive surgery, organ conservation, and restoration of function. Proceedings of the sixth course in urological oncology. Sicily, Italy, March 25-31, 1990." Prog Clin Biol Res **370**: 1-412.

Alvarez, C. P., F. Lasala, et al. (2002). "C-type lectins DC-SIGN and L-SIGN mediate cellular entry by Ebola virus in cis and in trans." J Virol **76**(13): 6841-6844.

Amundadottir, L. T., P. Sulem, et al. (2006). "A common variant associated with prostate cancer in European and African populations." Nat Genet **38**(6): 652-658.

Barreiro, L. B., E. Patin, et al. (2005). "The heritage of pathogen pressures and ancient demography in the human innate-immunity CD209/CD209L region." Am J Hum Genet **77**(5): 869-886.

Bollen, J., A. Polstra, et al. (2003). "Multicentric Castleman's disease and Kaposi's sarcoma in a cyclosporin treated, HIV-1 negative patient: case report." BMC Blood Disord **3**(1): 3.

Boyd-Patrick HA, P. A. (1997). "Tobago Health Needs Assessment ".

Bratt, O. (2002). "Hereditary prostate cancer: clinical aspects." J Urol **168**(3): 906-913.

Bratt, O., J. E. Damber, et al. (2002). "Hereditary prostate cancer: clinical characteristics and survival." J Urol **167**(6): 2423-2426.

Bunker, C. H., A. L. Patrick, et al. (2002). "High prevalence of screening-detected prostate cancer among Afro-Caribbeans: the Tobago Prostate Cancer Survey." Cancer Epidemiol Biomarkers Prev **11**(8): 726-729.

Bunker, C. H., A. L. Patrick, et al. (2002). "Prostate cancer risk is three-fold higher among men, aged 50-64, of African descent compared with men of Asian-Indian descent in Trinidad and Tobago." Ethn Dis **12**(4): S3-30-33.

Carpten, J., N. Nupponen, et al. (2002). "Germline mutations in the ribonuclease L gene in families showing linkage with HPC1." Nat Genet **30**(2): 181-184.

Carter, B. S., T. H. Beaty, et al. (1992). "Mendelian inheritance of familial prostate cancer." Proc Natl Acad Sci U S A **89**(8): 3367-3371.

Casey, G., P. J. Neville, et al. (2002). "RNASEL Arg462Gln variant is implicated in up to 13% of prostate cancer cases." Nat Genet **32**(4): 581-583.

Central Statistical Office of Tobago, R. o. T. (1993). 1990 Population and Housing Census. Demographic Report:
Age Structure, Religion, Ethnic Group, Education. **11:** 4-5.

Chang, B. L., S. L. Zheng, et al. (2002). "Polymorphic GGC repeats in the androgen receptor gene are associated with hereditary and sporadic prostate cancer risk." Hum Genet **110**(2): 122-129.

Chang, Y., E. Cesarman, et al. (1994). "Identification of herpesvirus-like DNA sequences in AIDS-associated Kaposi's sarcoma." Science **266**(5192): 1865-1869.

Chen, X., L. Levine, et al. (1999). "Fluorescence polarization in homogeneous nucleic acid analysis." Genome Res **9**(5): 492-498.

Cicek, M. S., X. Liu, et al. (2005). "Role of androgen metabolism genes CYP1B1, PSA/KLK3, and CYP11alpha in prostate cancer risk and aggressiveness." Cancer Epidemiol Biomarkers Prev **14**(9): 2173-2177.

Cramer, S. D., B. L. Chang, et al. (2003). "Association between genetic polymorphisms in the prostate-specific antigen gene promoter and serum prostate-specific antigen levels." J Natl Cancer Inst **95**(14): 1044-1053.

Crawford, E. D. (2003). "Epidemiology of prostate cancer." Urology **62**(6 Suppl 1): 3-12.

Crawford, E. D. (2009). "Understanding the epidemiology, natural history, and key pathways involved in prostate cancer." Urology **73**(5 Suppl): S4-10.

Dean, F. B., S. Hosono, et al. (2002). "Comprehensive human genome amplification using multiple displacement amplification." Proc Natl Acad Sci U S A **99**(8): 5261-5266.

Despres, P., A. Sakuntabhai, et al. (2005). "[A variant in the CD209 (DC-SIGN) promoter is associated with severity of dengue disease]." Med Sci (Paris) **21**(11): 905-906.

Dong, B., S. Kim, et al. (2007). "An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors." Proc Natl Acad Sci U S A **104**(5): 1655-1660.

Eder, I. E., Z. Culig, et al. (2000). "Inhibition of LncaP prostate cancer cells by means of androgen receptor antisense oligonucleotides." Cancer Gene Ther **7**(7): 997-1007.

Edwards, A., H. A. Hammond, et al. (1992). "Genetic variation at five trimeric and tetrameric tandem repeat loci in four human population groups." Genomics **12**(2): 241-253.

Eeles, R. A., Z. Kote-Jarai, et al. (2008). "Multiple newly identified loci associated with prostate cancer susceptibility." Nat Genet **40**(3): 316-321.

Engering, A., T. B. Geijtenbeek, et al. (2002). "The dendritic cell-specific adhesion receptor DC-SIGN internalizes antigen for presentation to T cells." J Immunol **168**(5): 2118-2126.

Feinberg, H., Y. Guo, et al. (2005). "Extended neck regions stabilize tetramers of the receptors DC-SIGN and DC-SIGNR." J Biol Chem **280**(2): 1327-1335.

Freedman, M. L., C. A. Haiman, et al. (2006). "Admixture mapping identifies 8q24 as a prostate cancer risk locus in African-American men." Proc Natl Acad Sci U S A **103**(38): 14068-14073.

Frison, N., M. E. Taylor, et al. (2003). "Oligolysine-based oligosaccharide clusters: selective recognition and endocytosis by the mannose receptor and dendritic cell-specific intercellular adhesion molecule 3 (ICAM-3)-grabbing nonintegrin." J Biol Chem **278**(26): 23922-23929.

Geijtenbeek, T. B., D. J. Krooshoop, et al. (2000). "DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking." Nat Immunol **1**(4): 353-357.

Geijtenbeek, T. B., D. S. Kwon, et al. (2000). "DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells." Cell **100**(5): 587-597.

Geijtenbeek, T. B., R. Torensma, et al. (2000). "Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses." Cell **100**(5): 575-585.

Giovannucci, E., E. B. Rimm, et al. (1993). "A prospective study of dietary fat and risk of prostate cancer." J Natl Cancer Inst **85**(19): 1571-1579.

Giovannucci, E., M. J. Stampfer, et al. (1997). "The CAG repeat within the androgen receptor gene and its relationship to prostate cancer." Proc Natl Acad Sci U S A **94**(7): 3320-3323.

Gsur, A., M. Preyer, et al. (2002). "Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk." Carcinogenesis **23**(10): 1647-1651.

Gudmundsson, J., P. Sulem, et al. (2007). "Genome-wide association study identifies a second prostate cancer susceptibility variant at 8q24." Nat Genet **39**(5): 631-637.

Gudmundsson, J., P. Sulem, et al. (2007). "Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes." Nat Genet **39**(8): 977-983.

Guo, Y., H. Feinberg, et al. (2004). "Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR." Nat Struct Mol Biol **11**(7): 591-598.

Hoffman, L. J., C. H. Bunker, et al. (2004). "Elevated seroprevalence of human herpesvirus 8 among men with prostate cancer." J Infect Dis **189**(1): 15-20.

Jemal, A., R. Siegel, et al. (2008). "Cancer statistics, 2008." CA Cancer J Clin **58**(2): 71-96.

Jeronimo, C., G. Varzim, et al. (2002). "I105V polymorphism and promoter methylation of the GSTP1 gene in prostate adenocarcinoma." Cancer Epidemiol Biomarkers Prev **11**(5): 445-450.

Kwok, P. Y. (2002). "SNP genotyping with fluorescence polarization detection." Hum Mutat **19**(4): 315-323.

Lai, J., M. A. Kedda, et al. (2007). "PSA/KLK3 AREI promoter polymorphism alters androgen receptor binding and is associated with prostate cancer susceptibility." Carcinogenesis **28**(5): 1032-1039.

Li, J., R. J. Coates, et al. "Steroid 5-{alpha}-reductase Type 2 (SRD5a2) gene polymorphisms and risk of prostate cancer: a HuGE review." Am J Epidemiol **171**(1): 1-13.

Lindmark, F., B. A. Jonsson, et al. (2004). "Analysis of the macrophage scavenger receptor 1 gene in Swedish hereditary and sporadic prostate cancer." Prostate **59**(2): 132-140.

Makridakis, N. M., R. K. Ross, et al. (1999). "Association of mis-sense substitution in SRD5A2 gene with prostate cancer in African-American and Hispanic men in Los Angeles, USA." Lancet **354**(9183): 975-978.

Martin, M. P., M. M. Lederman, et al. (2004). "Association of DC-SIGN promoter polymorphism with increased risk for parenteral, but not mucosal, acquisition of human immunodeficiency virus type 1 infection." J Virol **78**(24): 14053-14056.

McDonald, A. C., C. C. Ragin, et al. (2009). "Human herpesvirus 8 seroprevalence among Tobago women and the role of sexual lifestyle behavior." J Med Virol **81**(2): 264-270.

Medeiros, R., A. Morais, et al. (2002). "Linkage between polymorphisms in the prostate specific antigen ARE1 gene region, prostate cancer risk, and circulating tumor cells." Prostate **53**(1): 88-94.

Miljkovic-Gacic, I., R. E. Ferrell, et al. (2005). "Estimates of African, European and Native American ancestry in Afro-Caribbean men on the island of Tobago." Hum Hered **60**(3): 129-133.

Montgomery, J. D. (2008). THE ASSOCIATION OF HUMAN HERPESVIRUS 8 AND A SINGLE NUCLEOTIDE POLYMORPHISM IN THE GP130 SIGNALING RECEPTOR IN PROSTATE CANCER. Department of Infect. Pittsburgh, University of Pittsburgh. **PhD**.

Moore, P. S. and Y. Chang (2001). "Molecular virology of Kaposi's sarcoma-associated herpesvirus." Philos Trans R Soc Lond B Biol Sci **356**(1408): 499-516.

Nelson, C. P., L. C. Kidd, et al. (2001). "Protection against 2-hydroxyamino-1-methyl-6-phenylimidazo[4,5-b]pyridine cytotoxicity and DNA adduct formation in human prostate by glutathione S-transferase P1." Cancer Res **61**(1): 103-109.

Ntais, C., A. Polycarpou, et al. (2003). "SRD5A2 gene polymorphisms and the risk of prostate cancer: a meta-analysis." Cancer Epidemiol Biomarkers Prev **12**(7): 618-624.

Palapattu, G. S., S. Sutcliffe, et al. (2005). "Prostate carcinogenesis and inflammation: emerging insights." Carcinogenesis **26**(7): 1170-1181.

Rappocciolo, G., F. J. Jenkins, et al. (2006). "DC-SIGN is a receptor for human herpesvirus 8 on dendritic cells and macrophages." J Immunol **176**(3): 1741-1749.

Rivas, C., A. E. Thlick, et al. (2001). "Kaposi's sarcoma-associated herpesvirus LANA2 is a B-cell-specific latent viral protein that inhibits p53." J Virol **75**(1): 429-438.

Sakuntabhai, A., C. Turbpaiboon, et al. (2005). "A variant in the CD209 promoter is associated with severity of dengue disease." Nat Genet **37**(5): 507-513.

Salinas, C. A., M. A. Austin, et al. (2005). "Polymorphisms in the androgen receptor and the prostate-specific antigen genes and prostate cancer risk." Prostate **65**(1): 58-65.

Serth, J., F. Panitz, et al. (1999). "Increased levels of human papillomavirus type 16 DNA in a subset of prostate cancers." Cancer Res **59**(4): 823-825.

Severi, G., V. M. Hayes, et al. (2006). "Variants in the prostate-specific antigen (PSA) gene and prostate cancer risk, survival, and circulating PSA." Cancer Epidemiol Biomarkers Prev **15**(6): 1142-1147.

Snyder, G. A., J. Ford, et al. (2005). "Characterization of DC-SIGN/R interaction with human immunodeficiency virus type 1 gp120 and ICAM molecules favors the receptor's role as an antigen-capturing rather than an adhesion receptor." J Virol **79**(8): 4589-4598.

Soilleux, E. J., R. Barten, et al. (2000). "DC-SIGN; a related gene, DC-SIGNR; and CD23 form a cluster on 19p13." J Immunol **165**(6): 2937-2942.

Steinberg, G. D., B. S. Carter, et al. (1990). "Family history and the risk of prostate cancer." Prostate **17**(4): 337-347.

Stephens, M., N. J. Smith, et al. (2001). "A new statistical method for haplotype reconstruction from population data." Am J Hum Genet **68**(4): 978-989.

Tabarani, G., M. Thepaut, et al. (2009). "DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain." J Biol Chem **284**(32): 21229-21240.

Tavtigian, S. V., J. Simard, et al. (2001). "A candidate prostate cancer susceptibility gene at chromosome 17p." Nat Genet **27**(2): 172-180.

Thomas, G., K. B. Jacobs, et al. (2008). "Multiple loci identified in a genome-wide association study of prostate cancer." Nat Genet **40**(3): 310-315.

Tu, H., S. C. Jacobs, et al. (1994). "Rare incidence of human papillomavirus types 16 and 18 in primary and metastatic human prostate cancer." Urology **44**(5): 726-731.

Tut, T. G., F. J. Ghadessy, et al. (1997). "Long polyglutamine tracts in the androgen receptor are associated with reduced trans-activation, impaired sperm production, and male infertility." J Clin Endocrinol Metab **82**(11): 3777-3782.

Underhill, P. A., L. Jin, et al. (1997). "Detection of Numerous Y Chromosome Biallelic Polymorphisms by Denaturing High-Performance Liquid Chromatography." Genome Research **7**: 996--1005.

Urisman, A., R. J. Molinaro, et al. (2006). "Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant." PLoS Pathog **2**(3): e25.

Vastag, B. (2002). "Genome analysis yields mutations linked to hereditary prostate cancer." JAMA **287**(7): 827-828.

Wang, L., S. K. McDonnell, et al. (2002). "Analysis of the RNASEL gene in familial and sporadic prostate cancer." Am J Hum Genet **71**(1): 116-123.

Wang, L. Z., K. Sato, et al. (2003). "Polymorphisms in prostate-specific antigen (PSA) gene, risk of prostate cancer, and serum PSA levels in Japanese population." Cancer Lett **202**(1): 53-59.

Whittemore, A. S., L. N. Kolonel, et al. (1995). "Prostate cancer in relation to diet, physical activity, and body size in blacks, whites, and Asians in the United States and Canada." J Natl Cancer Inst **87**(9): 652-661.

Witte, J. S. (2009). "Prostate cancer genomics: towards a new understanding." Nat Rev Genet **10**(2): 77-82.

Wu, C. P. and F. L. Gu (1991). "The prostate in eunuchs." Prog Clin Biol Res **370**: 249-255.

Xu, J., S. L. Zheng, et al. (2003). "Common sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk." Am J Hum Genet **72**(1): 208-212.

Xu, J., S. L. Zheng, et al. (2002). "Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk." Nat Genet **32**(2): 321-325.

Xue, W., R. A. Irvine, et al. (2000). "Susceptibility to prostate cancer: interaction between genotypes at the androgen receptor and prostate-specific antigen loci." Cancer Res **60**(4): 839-841.

Xue, W. M., G. A. Coetzee, et al. (2001). "Genetic determinants of serum prostate-specific antigen levels in healthy men from a multiethnic cohort." Cancer Epidemiol Biomarkers Prev **10**(6): 575-579.

Yeager, M., N. Orr, et al. (2007). "Genome-wide association study of prostate cancer identifies a second risk locus at 8q24." Nat Genet **39**(5): 645-649.